INDEX

499-736—73——1

CITATIONS

In the Supreme Court of the United States

OCTOBER TERM, 1972

No. 72-414

HYNSON, WESTCOTT & DUNNING, INCORPORATED,
CROSS-PETITIONER

v.

CASPAR W. WEINBERGER, SECRETARY OF HEALTH, EDUCATION, AND WELFARE, AND CHARLES C. EDWARDS, COMMISSIONER OF FOOD AND DRUGS

ON WRIT OF CERTIORARI TO THE UNITED STATES COURT OF APPEALS FOR THE FOURTH CIRCUIT

BRIEF FOR THE RESPONDENTS

OPINIONS BELOW

The opinion of the court of appeals (J.A. 173–180) is reported at 461 F.2d 215. The order of the Commissioner of Food and Drugs was published in the Federal Register on June 18, 1971 (J.A. 72–78; 36 Fed. Reg. 11763).

JURISDICTION

The judgment of the court of appeals (J.A. 181) was entered on May 24, 1972. The petition for a writ of certiorari was filed, pursuant to an extension of

time granted by Mr. Justice Rehnquist, on September 11, 1972. On January 8, 1973, this Court granted the petition and consolidated the case with four other cases in which it also granted writs of certiorari.[1] The jurisdiction of this Court rests on 28 U.S.C. 1254(1) and 21 U.S.C. 355(h).

QUESTIONS PRESENTED

1. Whether, in a proceeding under Section 505(e) of the Federal Food, Drug, and Cosmetic Act to withdraw approval of a new drug application, the Commission of Food and Drugs has authority to determine initially whether the drug involved is a "new drug" within the meaning of Section 201(p) of the Act.

2. Whether a drug may be held not to be a "new drug" under Section 201(p)—that is, to be generally recognized among qualified experts as safe and effective for its intended use—in the absence of published adequate and well-controlled investigations of the type defined by Section 505(d) of the Act and the Commissioner's regulations.

3. Whether a drug for which a new drug application was on file with the Food and Drug Administration at the time of enactment of the Drug Amendments of 1962, which was at that time no longer a "new drug" under the definition contained in the 1938 Act, is exempted from the drug effectiveness requirements of the 1962 Amendments by Section 107(c)(4)

[1] The other cases are *Weinberger* v. *Hynson, Westcott and Dunning, Inc.*, No. 72–394, *CIBA Corporation* v. *Weinberger*, No. 72–528; *Weinberger* v. *Bentex Pharmaceuticals, Inc.*, No. 72–555; and *USV Pharmaceutical Corporation* v. *Weinberger*, No. 72–666.

thereof (and whether the application for such drug was not "effective" within the meaning of Section 107(c)(2) of the 1962 Amendments, so that approval of the application could not be withdrawn by the Food and Drug Administration).

STATUTORY PROVISIONS AND REGULATIONS INVOLVED

Section 201(p) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 321(p), J.A. 475) provides:

> The term "new drug" means—
>
> (1) Any drug (except a new animal drug or an animal feed bearing or containing a new animal drug) the composition of which is such that such drug is not generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use under the conditions prescribed, recommended, or suggested in the labeling thereof, except that such a drug not so recognized shall not be deemed to be a "new drug" if at any time prior to the enactment of this Act it was subject to the Food and Drugs Act of June 30, 1906, as amended, and if at such time its labeling contained the same representations concerning the conditions of its use; or
>
> (2) Any drug (except a new animal drug or an animal feed bearing or containing a new animal drug) the composition of which is such that such drug, as a result of investigations to determine its safety and effectiveness for use under such conditions, has become so recognized, but which has not, otherwise than in such investigations, been used to a material extent or for a material time under such conditions.

Section 107(c)(2) of the Drug Amendments of 1962, 76 Stat. 788, Note following 21 U.S.C. 321 (1970 ed.) (J.A. 481), provides:

> An application filed pursuant to section 505(b) of the basic Act which was "effective" within the meaning of that Act on the day immediately preceding the enactment date shall be deemed, as of the enactment date, to be an application "approved" by the Secretary within the meaning of the basic Act as amended by this Act.

Section 107(c)(4) of the Drug Amendments of 1962, 76 Stat. 789, Note following 21 U.S.C. 321 (1970 ed.) (J.A. 482), provides:

> In the case of any drug which, on the first day immediately preceding the enactment date, (A) was commercially used or sold in the United States, (B) was not a new drug as defined by section 201(p) of the basic Act as then in force, and (C) was not covered by an effective application under section 505 of that Act, the amendments to section 201(p) made by this Act shall not apply to such drug when intended solely for use under conditions prescribed, recommended, or suggested in labeling with respect to such drug on that day.

Other pertinent provisions of statutes and regulations are set forth at J.A. 475–491.

STATEMENT

This is a cross-petition to review the same decision of the court of appeals from which the government has petitioned in *Weinberger* v. *Hynson, Westcott and Dunning, Inc.*, No. 72–394. We respectfully refer the

Court to our brief in No. 72–394 (pp. 3–10) for a full statement of the case, and we recite here only the facts relevant to the issues addressed in this brief.

On October 16, 1970, a month after it lost its district court suit against the Commissioner for declaratory relief (J.A. 23–24), Hynson renewed its request for an administrative hearing (J.A. 24). Hynson asked the Commissioner to decide: (1) whether Lutrexin was, because of alleged current general recognition by experts of safety and effectiveness, not a "new drug" within the meaning of Section 201(p) of the Act; and (2) whether Lutrexin was, because of alleged general recognition of safety on October 9, 1962, "grandfathered" under Section 107(c)(4) of the 1962 Drug Amendments (J.A. 26).[2] In support of its request for a hearing on these issues, Hynson submitted certain affidavits of physicians and reports of medical studies which it had presented to the district court in its unsuccessful suit (J.A. 30–72, 87–121, 131–135).

On May 31, 1971, the Commissioner issued an order denying Hynson's request for a hearing and withdrawing approval of the new drug application ("NDA") for Lutrexin (J.A. 72–78). The Commissioner held that Lutrexin was not exempt under Section 107(c)(4) because the drug's NDA, which had become effective in 1953, had not been withdrawn prior to the enactment date of the 1962 Amendments, and thus the drug was

[2] Hynson also asked the Commissioner to determine whether under Section 505 of the Act there was substantial evidence of Lutrexin's effectiveness (J.A. 26).

"covered by an effective application" within the meaning of Section 107(c)(4)(C) (J.A. 74).[3] The Commissioner also ruled that Lutrexin was currently a new drug because there was no general recognition among experts of its effectiveness. The basis for this ruling was that Hynson had failed to present evidence of adequate and well-controlled clinical investigations in support of the drug's efficacy. In a detailed review of Hynson's documentation, the Commissioner demonstrated how each study showed deficiencies on its face which prevented it from being regarded as an adequate and well-controlled study (J.A. 74–78). He concluded, accordingly, that "there is no data base upon which experts can fairly and responsibly conclude that the safety and effectiveness of the drugs has been proven and is so well established that the drugs can be generally recognized among such experts as safe and effective for their intended uses" (J.A. 74).[4]

On petition for review, the court of appeals affirmed the Commissioner's ruling that Lutrexin is not exempt under Section 107(c)(4), although it held that the issue should have been decided by the district court and not by the Commissioner (J.A. 176–177). The

---

[3] The Commissioner did not take issue with Hynson's contention that on October 9, 1962, Lutrexin was generally recognized as safe by experts, that is, was not a new drug, and we assume for the purposes of this brief, therefore, that on that date Lutrexin was not a new drug as defined by the 1938 Act.

[4] The Commissioner also ruled that Hynson was not entitled to a hearing on the question of substantial evidence of effectiveness because it had failed to present adequate and well-controlled clinical investigations in support of its claim that substantial evidence of effectiveness exists (J.A. 78).

court neither affirmed nor rejected the Commissioner's holding that Lutrexin was currently a new drug.[5]

ARGUMENT

I. INTRODUCTION AND SUMMARY

The court of appeals implicitly held in this case that only the district court had authority initially to determine whether Lutrexin is a new drug (J.A. 176–177).[6] Hynson contends that the Commissioner has authority to determine new drug status in proceedings to withdraw approval of the product's new drug application ("NDA") under Section 505(e) of the Act (Br. 15–20). We agree. Since we have treated this issue at length in our briefs in *CIBA* (No. 72–528) and *Bentex* (No. 72–555), and since we do not disagree with the arguments advanced by Hynson on this point, we shall not discuss the issue further here.

---

[5] The court reversed the Commissioner's withdrawal of approval of the NDA for Lutrexin for failure to provide an opportunity for hearing on the question whether Hynson had presented substantial evidence of the drug's effectiveness (J.A. 181). The correctness of the court's ruling on this point is at issue in No. 72–394.

[6] The court explicitly held only that Hynson's claim for an exemption under Section 107(c)(4) was for the district court and not the Commissioner to decide (J.A. 176–177). However, since Section 107(c)(4) provides exemption from the definition of new drug contained in Section 201(p), and since in both its *Bentex* and *USV* opinions the court held that the district courts alone may decide whether a drug is a new drug (J.A. 266–267, 467), there is no doubt that the court here too meant that the Commissioner is without authority to decide whether Lutrexin is a new drug. We contend (see point II, *infra*) that the Commissioner correctly ruled that Lutrexin is a new drug, and we believe the court of appeals should have affirmed that ruling.

For the reasons stated in points II and III, *infra*, we do not believe that the existence of such jurisdiction aids Hynson in the present case, since it becomes pertinent only if this Court disagrees with the Commissioner's reasons for rejecting these claims. Hynson contends (1) that it is entitled to establish general recognition of safety and effectiveness under Section 201(p) by means of evidence that would not suffice to satisfy the requirements of Section 505 for approval of an NDA; and (2) that it can establish entitlement to exemption, pursuant to Section 107(c)(4), from the drug effectiveness requirements added to Sections 201(p) and 505 by the 1962 Amendments by demonstrating that on October 9, 1962, its product was no longer a "new drug" under the pre-Amendments definition.[7] If, contrary to our position, either of these contentions is accepted, we would agree that the Commissioner's action withdrawing approval of Lutrexin's NDA was based on erroneous premises, and that the matter should be remanded to the agency for further consideration—and that Hynson is entitled to a hearing before the agency on its claim for exemption, provided it can show that it has evidence that would support its claim under the standard established by the Court.

We show in point II, *infra*, that if there is a lack of substantial evidence of Lutrexin's effectiveness, so that approval of its NDA must be withdrawn pursuant to Section 505(e), that determination also serves

[7] Hynson's arguments in support of its right to a hearing on these issues (Br. 28–38) depend, of course, on acceptance of these contentions.

to foreclose the possibility that Lutrexin is "generally recognized" by qualified experts as effective. In adopting the requirement in the 1962 Amendments that drug effectiveness claims be supported by evidence consisting of "adequate and well-controlled investigations, including clinical investigations * * *," Congress heeded the unanimous testimony of a number of the Nation's leading clinical pharmacologists and medical educators that only such investigations, and not the clinical impressions of practicing physicians or poorly controlled studies, can provide a scientifically sound basis for the drug effectiveness inquiry. At the same time that Congress was amending Section 505 to incorporate the requirement of modern, scientific standards of drug efficacy evaluation, it also amended Section 201(p) to expand the definition of "new drug" to include drugs not generally recognized by experts as effective. Surely Congress did not intend to permit drugs to remain on the market, in the absence of adequate and well-controlled investigations supporting their efficacy claims, on the basis of "expert" opinions rooted in the very kind of unscientific evidence Congress rejected in Section 505.

The structure of the Act, in its normal application, confirms that the requirements of "general recognition" and material use contained in Section 201(p), which must be met in order for a drug not to be "new," are additional to the safety and effectiveness showing that must be made under Section 505. This is because there is no way for a drug to achieve general recognition and material use unless it has in fact

been widely written about and used; since distribution of the drug without an approved application is illegal, it follows that the drug must be able to satisfy the standards of Section 505 to obtain the approved application that will permit it ultimately to achieve "not new" status.

Lutrexin and other similarly situated drugs have escaped this normal situation because they were first marketed at a time when effectiveness was not a criterion under either Section 201(p) or Section 505. Nevertheless, it is clear that Congress intended such drugs, if not grandfathered under the 1938 Act or exempted by Section 107(c)(4) of the 1962 Amendments, to be subject to review for effectiveness and removal from the market for failure to satisfy the "substantial evidence" standards of Section 505. Congress specifically provided in Section 107(c)(3)(B) that applications with respect to drugs then on the market would, after a two-year grace period, be subject to withdrawal of approval under Section 505(e)(3). That provision specifies that approval shall be withdrawn if there is a lack of "substantial evidence" of effectiveness, which, of course, turns upon the existence of adequate and well-controlled investigations. Where it has been validly determined in administrative proceedings, in accordance with this congressional scheme, that there is a lack of substantial evidence of a drug's effectiveness, Congress surely did not intend to permit such a determination to be mooted by reliance upon insubstantial evidence of effectiveness, *i.e.,* opinions predicated on scientifically unreliable bases.

Hynson contends, alternatively, that Lutrexin is exempted from efficacy regulation by Section 107(c)(4) because it was, on the day preceding enactment of the 1962 Amendments, not a "new drug" under the safety-only standards then prevailing. It argues that, not being a new drug, its product was no longer subject to regulation under Section 505 and its application was therefore no longer "effective" within the meaning of Clause (C) of Section 107(c)(4). By thus linking the concept of an effective application to new drug status, however, Hynson would render Clause (C) wholly superfluous, since the requirements for exemption would be met by a drug satisfying Clauses (A) and (B). The construction adopted by the court of appeals, which recognizes that an application that had become effective prior to 1962 remained "effective" for purposes of Section 107(c)(4) (unless suspended by FDA) regardless of the subsequent history of the drug, gives a reasonable meaning to the provision and comports with the basic thrust of the 1962 Amendments to remove ineffective drugs from the market. That construction also comports with the unequivocal and repeated statements in the committee reports and the debates that the exemption in Section 107(c)(4) was intended to relate to drugs that had *never* been subject to new drug regulation under the pre-1962 standards—a condition Lutrexin obviously cannot satisfy.

The history of the enactment of the 1962 Amendments confirms the congressional intent to make drugs subject to re-evaluation on effectiveness grounds even

though they were not "new drugs" under the pre-1962 standard. The bill, as originally reported by the Senate Committee in July, 1962, would have exempted such drugs, but this initial proposal was significantly revised and tightened after the thalidomide tragedy, including the adoption of transitional provisions in Section 107(c) specifically providing for efficacy review of pre-1962 drugs, with the limited exception carved out by Section 107(c)(4). Under Hynson's interpretation, however, that exception would apply to nearly all drugs then on the market, while the regulatory provisions adopted by Congress would apply only to the relatively small group which, although approved on safety grounds by FDA, were not yet generally recognized as safe. There is no warrant for such a construction, which would almost entirely defeat the regulatory thrust of the transitional provisions.

Finally, it is clear from the structure of the original Act that applications in fact remained "effective" in a significant sense (although perhaps inactive) even after a drug had ceased to be "new." This is because Section 505(e) directed suspension of the effectiveness of an application in the event new information came to light questioning the drug's safety.

II. THE EXISTENCE OF PUBLISHED ADEQUATE AND WELL-CONTROLLED CLINICAL INVESTIGATIONS IS A PREREQUISITE TO A FINDING UNDER SECTION 201(p) OF GENERAL RECOGNITION OF A DRUG'S EFFECTIVENESS

The Commissioner ruled that Lutrexin is a "new drug" as defined in Section 201(p) of the Act. He stated (J.A. 74):

> No adequate and well-controlled clinical investigations published in the medical literature have been identified. Therefore, there is no data base upon which experts can fairly and responsibly conclude that the safety and effectiveness of the drugs has been proven and is so well established that the drugs can be generally recognized among such experts as safe and effective for their intended uses.

Hynson contends that the Commissioner applied the wrong standard to the "new drug" inquiry. While it concedes that the existence of adequate and well-controlled clinicial investigations is essential to a finding of a new drug's effectiveness under Section 505, it nonetheless argues that status as a "new drug" under Section 201(p), which turns upon "general recognition" among qualified experts of a drug's effectiveness, does not require the existence and publication of scientifically valid evidence of effectiveness (Br. 21–27). Our view, on the other hand, is that the statutory scheme and the history underlying the adoption by Congress of scientific criteria for evaluation of drug effectiveness show that Congress acted on the premise that "experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs," whose collective judgment is the subject of the Section 201(p) inquiry, would necessarily predicate that judgment on the kind of evidence—or lack of it—required by Section 505; by the same token, the scientific community would necessarily reject, as a basis for general recognition of effectiveness, the kind of uncontrolled clinical impressions upon

which Hynson seeks to rely, unless corroborated by adequate and well-controlled clinical investigations.

Section 201(p) defines the term "new drug" for purposes of the Act. Under this definition, a drug is a "new drug" unless (1) it is generally recognized among qualified experts as safe and effective for its intended use, *and* (2) it has been so used to a material extent or for a material time (see Section 201(p)(1) and 201(p)(2) at J.A. 475). The Act, however, nowhere specifically defines what constitutes general recognition among qualified experts. Therefore, in order to give content to that term, and thus to the definition of "new drug," it is appropriate to look to the purposes to be achieved by the Act, as reflected in the related provisions of the Act and in its legislative history. Cf. *United States* v. *Bacto-Unidisk*, 394 U.S. 784.

In our view, the most pertinent related provision of the 1962 Amendments is the definition of "substantial evidence" of effectiveness in Section 505(d). Our contention, in essence, is that the general recognition among qualified experts to which Congress referred in Section 201(p) should be interpreted harmoniously with the definition of "substantial evidence" adopted by Congress in Section 505(d), and thus should not be construed to include any concept of "recognition" based on less than "substantial evidence" of effectiveness. Hynson contends the contrary. Our position on this issue is, we submit, more consistent both with the legislative history of the 1962 Amendments and with the overriding purpose of the

Amendments to insure that drugs allowed upon the market are effective for their claimed uses as well as safe. Cf. *United States* v. *Bacto-Unidisk, supra,* 394 U.S. at 798; *United States* v. *Sullivan,* 332 U.S. 689, 693–695; *United States* v. *Dotterweich,* 320 U.S. 277, 280.

A. THE HISTORY OF THE 1962 DRUG AMENDMENTS SHOWS THAT CONGRESS WAS PERSUADED THAT ONLY ADEQUATE AND WELL-CONTROLLED CLINICAL INVESTIGATIONS COULD PROVIDE A SCIENTIFICALLY RELIABLE BASIS FOR EVALUATION OF DRUG EFFECTIVENESS CLAIMS

The history of the Drug Amendments of 1962 leaves no doubt that Congress intended a rigorous pre-marketing examination of drugs to assure that a substantial scientific basis exists for their claims of effectiveness, and that it further intended that, even after approval and marketing, drugs would continue to be under supervision for safety and effectiveness thereafter, in light of evolving medical knowledge, so that products disclosed by new data to be unsafe or ineffective would not remain on the market. The system of pre-marketing clearance and post-marketing supervision by the expert administrative agency was the primary means established for exclusion of ineffective products from the market.

In Section 505(d), as amended in 1962, Congress adopted a definition of the kind of evidence to be considered "substantial evidence" of drug effectiveness claims. That definition reflects modern, scientific concepts of pharmacology. The kind of evidence required consists of "adequate and well-controlled investigations, including clinical investigations, by ex-

perts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have * * *'' (21 U.S.C. 355(d), J.A. 478–479).

This congressional insistence upon substantiation of drug efficacy claims by means of rigorous, scientific evidence of a kind and quality that would be acceptable to expert pharmacologists—rather than by reliance upon the clinical impressions of practicing physicians or the results of inadequately designed and poorly controlled experiments—was the product of careful study and of the presentation of uncontradicted testimony in congressional hearings by many of the Nation's leading experts on clinical pharmacology, by prominent medical educators, and by representatives of the drug industry itself.[8] Appendix A, pp. 1a–12a, *infra,* outlines the testimony on this subject presented at the congressional hearings.

The experts explained to Congress in detail the need for controlled clinical testing and the difference between the validity of such scientific procedures and the unreliability of clinical impressions of therapeutic effectiveness by individual practitioners. The testimony drew a sharp distinction between the expert qualified to evaluate the effectiveness of drugs—named in both Sections 201(p) and 505(d)—and the working physician who applies to individual patients

[8] Drug Industry Antitrust Act, Hearings before the Subcommittee on Antitrust and Monopoly, Senate Judiciary Committee, 87th Cong., 1st Sess, pursuant to S. Res. 52 on S. 1552.

contemporary medical knowledge developed by expert scientific research. As one leading pharmacologist testified, the "'average practicing physician—and I haved helped to train hundreds of them—just does not have the time, the facilities, the skill, nor the training to be an expert in the determination of drug efficacy.'"[9] This testimony was echoed by the other expert witnesses, who in substance agreed with the observation of one of their number that "[t]he history of medicine is, unhappily, replete with examples of useless drugs employed for years, decades or centuries, by countless physicians before a few properly conducted experiments proved the drugs to be without value."[10] Confirming this view, the Chairman of the Pharmaceutical Manufacturers Association emphasized that in determining effectiveness the key words are "'well-conducted clinical trials by competent clinicians.'"[11]

In addition, Congress was informed in detail of the meaning of an adequate and well-controlled study. Such a study is designed to eliminate subjective influence on the part of either the investigator or the subject; it is conducted on a comparative basis, with a sufficient number of subjects to minimize the influence of chance factors on results; it is of sufficient duration to permit all consequences of the drug's administration to be evaluated; and its results can be confirmed by other clinical investigators using similar

[9] App A, *infra*, p. 4a.
[10] *Id.* at 5a.
[11] *Id.* at 8a.

methods.[12] The views expressed by these witnesses to Congress represent the consensus of experts in the evaluation of drug effectiveness. The reasons for these rigorous scientific requirements are explained more fully in Appendix B, pp. 13a–47a, *infra*.

Thus, in Section 505(d) Congress rejected as unreliable determinations of drug effectiveness based upon clinical impressions or "majority vote" of practitioners,[13] and adopted instead the criterion of adequate and well-controlled clinical investigations. Hynson nevertheless contends that for purposes of Section 201(p) "a drug may have enjoyed such widespread use and therapeutic success that qualified experts could conclude that it is generally recognized as safe and effective, even if no studies of any kind existed in the medical literature" (Br. 27). Under this view, "experts qualified * * * to evaluate the * * * effectiveness of drugs" may be argued to have based a determination of effectiveness solely upon the clinical impressions of individual practitioners and the existence of scientifically inadequate and poorly controlled studies—the very criteria Congress rejected as scientifically unreliable by enacting Section 505(d).[14]

We submit, instead, that Section 201(p) should be construed harmoniously with Section 505(d) and the

---

[12] *Id.* at 2a–3a.

[13] *Id.* at 12a.

[14] It is particularly significant for this purpose that Section 505(d)'s definition of "substantial evidence" is in terms also applicable to proceedings under Section 505(e) for the withdrawal of an NDA, which can of course occur after the kind of widespread use to which Hynson refers.

congressional judgment the latter provision represents, by interpreting "general recognition" of effectiveness to mean more than, rather than less than, "substantial evidence" of effectiveness. The logic of this interpretation becomes apparent when the application of the two provisions in the normal, entirely prospective operation of the 1962 Amendments is considered.

B. THE NORMAL OPERATION OF THE REGULATORY SCHEME EMBODIED IN THE ACT REQUIRES DRUGS TO SATISFY THE CRITERIA FOR NDA APPROVAL IN SECTION 505 BEFORE THEY CAN CEASE TO BE "NEW DRUGS" UNDER THE DEFINITION IN SECTION 201(p)

In the ordinary course of entirely prospective operation of the 1962 Amendments, a drug could not acquire the characteristics necessary to cease being a "new drug" as defined in Section 201(p) unless it has first satisfied the safety and effectiveness standards of Section 505. The terms of Section 201(p) preclude any other result, since the provision renders it impossible to cease being a new drug without being used—Section 201(p)(2) explicitly requires use "to a material extent or for a material time," in addition to general recognition of effectiveness, in order for a drug to cease being a "new drug" (J.A. 475).

Since loss of "new drug" status can thus occur only after use, and since distribution of a "new drug" is unlawful without an application approved pursuant to the standards of Section 505, which require the manufacturer to produce "substantial evidence" of effectiveness in the form of "adequate and well-controlled investigations," the Act is designed so that

drugs on the market, unless exempt, will have mustered the requisite scientifically reliable evidence of effectiveness long before they are in a position to drop out of active regulation by ceasing to be a "new drug."

The substantial evidence of effectiveness initially offered to the Commissioner to obtain a marketing authorization does not, of course, establish that the drug is generally recognized as effective. It reflects only the scientific conclusions of the particular experts who performed the investigations that the product is therapeutically useful and their expectation that their fellow experts will agree with their conclusions when the scientific basis for those conclusions is made known to them. Thus, "not new drug" status comes about after these studies, or subsequent studies of the same kind, are published in the medical literature and inform other experts qualified in evaluating drugs of the product's effectiveness. Such studies then form the basis for general recognition by the experts that the drug is effective for its claimed uses. Once the product attains a consensus of expert acceptance, plus the actual usage of the product for a material time or to a material extent required by Section 201(p)(2) (primarily to provide greater insurance of safety), it may be viewed as currently exempt from Section 505's requirements, although its "new drug" status may be revived if expert knowledge changes.[15]

It is thus apparent that the statutory scheme is designed to condition "not new drug" status on the ability to satisfy the safety and effectiveness criteria

[15] See our brief in *USV* (No. 72–666) at p. 48, n. 31.

of Section 505, plus the ingredients added by Section 201(p)—general recognition of that ability plus material use.[16] Hynson seeks to fashion a loophole in this statutory structure by virtue of the fact that Lutrexin came on the market at a time when Section 505 contained different (safety only) standards for pre-marketing clearance than prevail today. But it is clear from the applicability and purpose of the 1962 Amendments, as discussed immediately below, that Congress did not intend products such as Lutrexin to escape scientific evaluation of their effectiveness.

C. THE 1962 AMENDMENTS CLEARLY CONTEMPLATED THAT DRUGS THEN ON THE MARKET, IF NOT EXPRESSLY EXEMPTED, WOULD BE SUBJECT TO REMOVAL FROM THE MARKET FOR FAILURE TO SATISFY THE DRUG EFFECTIVENESS STANDARDS OF SECTION 505

As is apparent from the structure and legislative history of the 1962 Amendments (see our brief in *USV,* pp. 17–29), Congress decided to subject all pre-1962 drugs to efficacy regulation except those grandfathered under the 1938 Act[17] and those specifically exempted by Section 107(c)(4). We show in point III, *infra,* that Lutrexin is not entitled to any such exemption from efficacy regulation. Accordingly, Sec-

[16] This view is further corroborated by the very terms of Section 201(p)(2), which indicate that Congress contemplated that general recognition of effectiveness would occur "as a result of investigations" and that such recognition could occur in the absence of material use of the drug.

[17] The grandfather provision contained in Section 201(p)(1) covers drugs subject to the 1906 Act, provided their labeling remains unchanged. Developments in pharmacology in the last 35 years have resulted in vast changes in products and claims, so that this exemption is of diminishing importance.

tion 107(c)(3)(B) (J.A. 482) makes it clear that, following the two-year grace period, approval of Lutrexin's NDA was subject to withdrawal under "clause (3) of the first sentence of section 505(e) * * * as amended * * *." That clause specifically provides for withdrawal of approval if "there is a lack of substantial evidence" that the drug is effective for its claimed uses.

In short, assuming we are correct that neither Section 107(c)(2) nor Section 107(c)(4) exempts drugs such as Lutrexin from administrative review for efficacy, Congress deliberately specified that such drugs should be evaluated for effectiveness on the basis of the existence of adequate and well-controlled clinical investigations. It is hardly conceivable that Congress intended that, after the Commissioner had made such an evaluation in the manner and pursuant to the standards Congress established and had found a product wanting, that product should nevertheless be able to stay on the market by producing the kind of evidence Congress had rejected as unscientific. A construction of the Act permitting such a result would be unsound, for it would "* * * impute to Congress a purpose to paralyze with one hand what it sought to promote with the other." *Clark* v. *Uebersee Finanz-Korp.*, 332 U.S. 480, 489.

Since the Act is not simply "a collection of English words," but a "working instrument of government" whose purposes are to protect the lives and health of the people (*United States* v. *Dotterweich*, 320 U.S. 277, 280), the construction urged by Hynson, under which the scientific standards of evaluation of drug

effectiveness selected by Congress may be discarded in favor of outmoded and unreliable criteria, should be rejected. Nothing in the statute requires that it be construed to enable a drug that had been examined under the criteria Congress prescribed and found wanting to remain on the market,[18] and the very purpose of the Act indicates that such a construction should be avoided.

[18] Hynson cites (Br. 32–35) a number of cases for the proposition that the determination under Section 201(p) is different from the determination under Section 505. While we disagree with the reasoning in some of these decisions, their holdings are not contrary to the government's argument here. All of these cases were proceedings involving products with respect to which no marketing authorization was on file. Thus the agency had never made a formal determination, in a proceeding under Section 505, whether the product met the safety and effectiveness standards for approval of an application. Moreover, with one possible exception, the courts in these cases refused in the absence of published reports of adequate and well-controlled studies to hold that the products involved were exempt. The possible exception is *United States* v. *Article of Drug Labeled "Quick-O-ver"*, 274 F. Supp. 443 (D. Md.), in which the effectiveness of the individual ingredients involved for the uses claimed was substantiated in the scientific literature and the court did not discuss this issue in its opinion. A more recent decision, *United States* v. *An Article of Drug* * * * *Xerac*, CCH Food, Drugs, Cosmetic L. Rep. ¶ 40,836 (N.D. Ill.), accepted the views expressed in this brief. And other recent authorities recognize that lack of documented sources of information bars the widespread expert knowledge of the drug's effectiveness or safety necessary to achieve "not new drug" status under Section 201(p). See *e.g.*, *United States* v. *An Art. of Drug* * * * *"Bentex Ulcerine"*, 469 F. 2d 875 (C.A. 5), petition for a writ of certiorari pending, No. 72–1198; *United States* v. *41 cases, more or less*, 420 F. 2d 1126, 1130– (C.A. 5); *United States* v. *Article of Drug (Norwich Pharmacal)*, 415 F. 2d 390 (C.A. 5), affirming, 294 F. Supp. 1307 (N.D. Ga.); *United States* v. *Article of Drug* * * * *"Mykocert"*, 345 F. Supp. 571, 574 (N.D. Ill.).

III. SECTION 107(c) OF THE 1962 AMENDMENTS DOES NOT EXEMPT LUTREXIN FROM WITHDRAWAL OF APPROVAL OF ITS NDA ON GROUNDS OF FAILURE TO ESTABLISH EFFECTIVENESS

Petitioner contends that Sections 107(c)(2) and 107(c)(4) of the 1962 Amendments exempt it from withdrawal of approval of its NDA on effectiveness grounds. This contention is predicated solely on the fact—which we concede *arguendo* for purposes of this case—that Lutrexin was no longer a "new drug" on October 9, 1962, because it was "generally recognized" as safe at that time. Petitioner reasons that, therefore, its NDA, which had become effective in 1953, had ceased to be "effective" by 1962, since at that time an NDA was no longer required for the marketing of drugs containing lututrin to be used as prescribed in Lutrexin's labeling.

We submit that the court of appeals was correct in rejecting this argument and in holding that an application that had become effective before October 10, 1962, remained "effective" for purposes of Section 107(c) regardless of the subsequent history of the drug (with the limited exception, not here relevant, of NDAs whose effectiveness had been suspended by FDA under old Section 505(e)[19]). This

[19] Petitioner suggests that this limited exception to the continuing effectiveness of an NDA leads to a "bizarre result" contrary to the purpose of Section 505(e) to prevent further distribution of unsafe or ineffective drugs (Br. 42–43). (See also USV's brief in No. 72–666, at p. 71.) Of course, the fact that the effectiveness of an NDA had been suspended by FDA for safety reasons would hardly qualify the product for Section 107(c)(4)'s exemption from efficacy regulation, since

conclusion is required both by the structure of the statute and by the clear congressional explanation of its intended operation.

A. LUTREXIN IS NOT ENTITLED TO EXEMPTION UNDER SECTION 107(C)(4)

Petitioner contends that Lutrexin is exempt under Section 107(c)(4) from the effectiveness requirements of the 1962 Amendments (Br. 41–48). Section 107(c)(4) provides that, so long as the manufacturer does not alter the drug's labeling, the addition of effectiveness to the defintion of "new drug" in Section 201(p) of the Act shall not apply to any drug which, on October 9, 1962: (A) was commercially used or sold in the United States; (B) was not a new drug as defined by the 1938 Act; *and* (C) was not covered by an effective application under Section 505 of that Act (J.A. 482). Petitioner contends that because a drug that was no longer a "new drug" had no further need of an effective application, it follows that such a drug's application was no longer "effective" within the meaning of Clause (C).

1. The first difficulty with petitioner's interpretation is, as the court of appeals stated (J.A. 176–177, citing J.A. 469–470), that it would render Clause (C)

---

such a product could not meet Clause (B)'s requirement that it be generally recognized as safe on October 9, 1962. There is nothing bizarre or inconsistent in the conclusion that any drug that had ever required an effective NDA to be marketed prior to 1962 was excluded from the exemption of Section 107(c)(4)—by Clause (B) in the case of drugs with suspended NDAs and by Clause (C) in the case of those with effective ones.

entirely superfluous, since the simple fact that Clause (B) is satisfied would suffice to meet the requirement of Clause (C). Plainly, however, Clause (C) is a requirement additional to that in Clause (B), since Congress employed the conjunctive "and" in listing the requirements that had to be satisfied for the exemption. Thus, only the interpretation adopted by the court of appeals—that all applications that became effective between 1938 and 1962 were "effective" for purposes of Clause (C), unless their effectiveness had been suspended by FDA—gives rational meaning to the provision.

2. Petitioner does not, in fact, propose any alternative construction of Clause (C) that would give it meaning. Instead, it seeks to defend its construction by claiming that the possible alternatives are equally defective, either because Clause (B) would be rendered surplusage or because the entire exemption provision would not be given significant meaning (Br. 43–44). Both of these suggestions are, however, plainly in error. Clause (B) is obviously necessary to insure that drugs illegally on the market because no one ever procured an effective NDA for them would not be granted exempt status. And the government's interpretation of Clause (C) does not "defeat the intention of Congress to exempt some drugs from the effectiveness provisions of the 1962 Amendments" (Br. 44). As we explain in our brief in the *USV* case (pp. 49–52), the provision was intended to and does exempt drugs that, as a generic class, were never subject to new drug regulation—a category consisting primarily

of over-the-counter drugs which, because of changed composition or labeling, were not "grandfathered" under the 1938 Act, but, because of universal recognition of the safety of their old, established ingredients at the time they came on the market, were never subject to new drug regulation.

3. Hynson also errs in suggesting that its new drug application could not "in logic nor law" have been effective once the drug it covered ceased to be a new drug (Br. 40). Certainly the 1938 Act provided no mechanism, other than the Commissioner's suspension authority under Section 505(e), by which an NDA, once it became effective, would cease to be effective thereafter. Indeed, the wording of the suspension provision itself clearly establishes that an application would not cease to be effective under the Act when the drug ceased to be "new." The very contingency Section 505(e) contemplates is that evolving scientific knowledge may cast doubt upon the safety of a drug previously considered safe,[20] and the provision is designed to give FDA power to deal administratively with such a contingency. If the NDA ceased to be effective because the drug it covered had (temporarily) ceased to be new, how could its "effectiveness" be suspended when doubt was cast on its safety? Only by treating the NDA as having continuing effectiveness could administrative control through Section 505's regulatory procedures be retained. The statute thus clearly contemplated that the NDA would remain effective in this sense even if

[20] See our brief in *USV* at p. 48, n. 31.

inactive, and it is petitioner's effort to shed this status as a means of escaping efficacy regulation for its product that is not supportable "in logic nor law."

4. Hynson also suggests that the legislative history of the 1962 Amendments supports its position that Congress meant to exempt drugs that had become generally recognized as safe from efficacy regulation (Br. 45–48). However, Hynson never reconciles this suggestion with the clear congressional explanation in the Senate, House, and Conference reports that the exemption afforded by Section 107(c)(4) is for drugs that had *never* been subject to new drug regulation. See S. Rep. No. 1744, Part 2, 87th Cong., 2d Sess., p. 8; H. Rep. No. 2464, 87th Cong., 2d Sess., p. 12; H. Rep. No. 2526, 87th Cong., 2d Sess., pp. 22–23; see also 108 Cong. Rec. 17366 (remarks of Senator Eastland). Lutrexin obviously fails to satisfy that congressional standard.

Moreover, to the extent inferences are to be drawn from the less explicitly applicable facets of the history of the 1962 Amendments, they show that Congress considered but ultimately refused to enact the wholesale exemption for which Hynson contends.

As originally reported by the Senate Judiciary Committee on July 19, 1962, the Amendments left the 1938 Act's definition of "new drug" in Section 201(p) unchanged. The new drug effectiveness provisions appeared solely in Section 505, to which was added the

requirement that, to obtain or retain an NDA, the manufacturer show substantial evidence of effectiveness. S. Rep. No. 1744, *supra,* Part 1, pp. 24–27. The effect of the bill in this version would have been that the manufacturer of a drug generally recognized as safe (*i.e.,* not a new drug) need not show that the drug is also effective. As the Senate report made clear, the bill was designed to protect such drugs in their current claims: "The Committee decided that it was unnecessary to add the requirement of 'effectiveness' to the definition of a new drug in section 201(p) * * *. If such a change were made in the definition of new drug, many 'old' established drugs might have to go through burdensome new drug clearance procedures even though their safety is unquestioned. That result is believed unsound." *Id.* at 17. Had this version of the bill become law, Lutrexin would indeed have been "grandfathered." However, it never became law.

Instead, President Kennedy, anxious to strengthen the bill, sent the Chairman of the Senate Judiciary Committee seven proposed amendments on August 4, 1962. 108 Cong. Rec. 15695–15698. The fourth proposed amendment called for the inclusion of "effectiveness" in Section 201(p)'s definition of "new drug" to "assure that all new drugs will have to be proved effective as well as safe for the uses for which they are offered * * *." *Id.* at 15696. The President also proposed a transition provision which would require "a commercially established drug to go through

the 'new drug' process" "if there is substantial doubt as to its efficacy." *Id.* at 15696.[21]

On August 21, 1962, the Senate Judiciary Committee reported a much strengthened version of the bill, with changes designed further to "insure the reliability of drugs." S. Rep. No. 1744, *supra,* Part 2, p. 1. In accordance with the President's suggestion, Section 201(p) now defined "new drug" to include effectiveness. *Id.* at 5. Sections 201(p) and 505 were thus made congruent in their use of a standard of effectiveness. Also in keeping with the President's suggestion, the Committee reported transitional provisions relating to drugs already on the market. These provisions, virtually identical to Section 107(c) as enacted, gave existing drugs which were generally recognized as safe but not generally recognized as effective a two-year grace period during which their NDAs were protected against withdrawal of approval for lack of substantial evidence of effectiveness. *Id.* at 7–8. The only type of drug which was exempted from the new requirements of effectiveness in Sec-

[21] The "substantial doubt" language gave way in the ultimate version of the legislation to the present language of "lack of substantial evidence." Thus, there was a modification in the standard by which the effectiveness of a drug was to be measured. But, contrary to Hynson's argument (Br. 45), this change in no way relieved pre-existing drugs of the need to prove effectiveness; it merely altered their burden or proof.

tions 201(p) and 505 was "* * * a new drug on the market which was never subject to the new drug procedure before * * *." *Id.* at 8.[22] The Senate then passed the bill with the effectiveness and transitional provisions as reported, and these were the provisions ultimately enacted.

This history thus establishes that Congress abandoned its original inclination to exempt "safe" pre-1962 drugs from the new drug effectiveness requirements, changing instead to a scheme under which all such drugs, unless grandfathered under the 1938 Act or coming within the narrowly defined exemption provided in Section 107(c)(4), were to be subject to efficacy regulation after the two-year grace period (presumably designed to afford manufacturers an opportunity to substantiate their products' claims by scientifically reliable investigations of the kind newly required in amended Section 505(d)).

---

[22] Petitioner contends (Br. 46) that because Part 2 of the Senate report made clear that a new or changed claim for an existing drug would require submission to the Commissioner of substantial evidence of effectiveness, the Senate did not intend to subject already existing claims to the effectiveness requirements. This is incorrect. The committee, in the same paragraph which petitioner cites, referred readers interested in the effect of the amendment on existing drugs to the discussion of the transitional provisions. S. Rep. No. 1744, *supra*, Part 2, p. 5. And, as we have shown above, those transitional provisions gave no permanent exemption to existing drugs which had been subject to the new drug procedures under the 1938 Act, whatever their claims.

B. SECTION 107(C)(2) DOES NOT EXEMPT HYNSON FROM ADMINISTRATIVE WITHDRAWAL OF APPROVAL OF LUTREXIN'S NDA ON EFFECTIVENESS GROUNDS

Hynson also contends (Br. 38–41) that its application was not "effective" within the meaning of Section 107(c)(2), so that it was not "deemed * * * approved" by that provision and was accordingly not subject to withdrawal of approval under Section 107(c)(3). There is no basis for supposing that the word "effective" was employed in any different sense in Section 107(c)(2) than it was in Section 107(c)(4)—and Hynson makes no such suggestion. Accordingly, much of the foregoing argument is equally applicable to this contention.

Section 107(c)(2) was included in the transitional provisions for the simple reason that the 1962 Amendments altered the prior regulatory procedures, under which applications became automatically effective (unless the Commissioner refused to permit them to become effective) after the passage of a prescribed period of time. Amended Section 505 requires affirmative approval of applications. By providing that effective applications would be "deemed * * * approved" (J.A. 481), Congress eliminated the necessity that every application then in effect be reviewed and approved; it also protected the marketing authority of manufacturers that had effective applications, at least until such time

as the approval conferred by Section 107(c)(2) might be withdrawn under the provisions of Section 107(c)(3). Section 107(c)(2) is thus nothing more than a housekeeping measure to smooth the transition from the old regulatory standards to the new ones. There is no evidence that it was intended to go further and confer any form of substantive immunity from regulation, as Hynson suggests.

In any event, unless Hynson's interpretation of the exemption provision in Section 107(c)(4) is accepted (in which case the interpretation of Section 107(c)(2) is of no consequence), it is difficult to see what Hynson can gain by its interpretation of Section 107(c)(2). If Hynson's NDA was not "deemed * * * approved," that simply puts it in the same posture it is now in after withdrawal of approval under Section 505(e)—that of marketing an unapproved drug. All that Hynson could succeed in establishing, therefore, is that its marketing of Lutrexin has been illegal since the enactment of the 1962 Amendments, because it would no longer be able to claim approved status under Sections 107(c)(2) and 107(c)(3), which status persists until administrative withdrawal of approval.

CONCLUSION

For the foregoing reasons and the reasons stated in our brief in No. 72-394, the judgment of the court of appeals should be reversed and the order of the Commissioner of Food and Drugs withdrawing approval

of the new drug application for Lutrexin should be affirmed.

Respectfully submitted.

ERWIN N. GRISWOLD,
Solictor General.
THOMAS E. KAUPER,
Assistant Attorney General.
LAWRENCE G. WALLACE,
Deputy Solicitor General.
ANDREW L. FREY,
Assistant to the Solicitor General.
HOWARD E. SHAPIRO,
ROBERT B. NICHOLSON,
Attorneys.

PETER BARTON HUTT,
Assistant General Counsel,
EUGENE PFEIFER,
Attorney,
Food, Drugs, and Product Safety Division, Department of Health, Education, and Welfare.

APRIL 1973.

APPENDIX A

Congressional Understanding of The Scientific Principles Underlying Adaquate and Well-Controlled Investigatons To Prove Drug Effectivness.

The scientific principles underlying adequate and well-controlled clinical investigations were described in detail to Congress during its consideration of the Drug Amendments of 1962 and formed the basis for the statutory requirements for proof of drug effectiveness. This Appendix sets forth excerpts from the large body of testimony of distinguished scientists, as well as representatives of the drug industry itself, urging upon Congress the importance of establishing scientifically valid criteria for evaluating drug effectiveness. Congress adopted these expert recommendations by incorporating the definition of "substantial evidence" of drug effectiveness in Section 505(d) of the Act.

1. On April 20, 1960, Dr. Austin Smith, President of the Pharmaceutical Manufacturers Association, appeared before the Senate Subcommittee on Antitrust and Monopoly. In the course of his testimony, an article co-authored by him was placed in the record: Smith et al., *Laboratory and Clinical Appraisals of New Drugs,* from the Journal of the American Medical Association, December 9, 1944.[1] While the authors acknowledged that clinical investigation of new drugs presents difficult

[1] Hearings on Administered Prices before the Subcommittee on Antitrust and Monopoly of the Senate Judiciary Committee, 86th Cong., 2d Sess., 10986–10991 (1960).

problems, they suggested that certain factors be considered:

(*a*) *The Selection of Individuals to be Observed.*—Cooperation from subject; absence of complicating factors; age; sex; emotional and psychic factors.

(*b*) *Diagnosis.*—Objective proof of diagnosis if possible, such as isolation and identification of infecting organism; x-ray evidence or other informative laboratory data; accurate description of lesion; differential diagnosis.

(*c*) *Control Observations.*—Preliminary control observations on the individuals; concurrent observations of untreated controls; alternation of treatment; alternation of treated and control subjects; posttreatment, control observations.

(*d*) *Observation During Treatment.*—Repeated physical and laboratory examinations; hematology; urinalysis; blood chemistry; x-ray; functional tests; precise objective measurements of improvements alleged to be produced by drug; determination of concentration of drug in blood, urine and other body fluids and the correlation of levels so observed with the dosage and the effect.

(*e*) *Number of Subjects.*—Sufficient number of treated individuals to minimize chance or other uncontrollable factors from influencing results; sufficient number of untreated control subjects. The results should be subjected to statistical analysis in order to determine their reliability.

(*f*) *Carefully Planned Administration of Drug.*—Controlled variation in dosage, frequency, method and duration of administration; effect of other drugs, and so on.

(*g*) *Criteria of Benefit.*—Establishment of criteria whereby the effects of the drug may be evaluated; objective tests; subjective observations; comparison with control treatments; comparison with natural course of disease.

(*h*) *Separation of Subjective and Objective Observations.*—Use of "blind tests" (neither investigator nor subject knows which of several samples being administered is control or test product) or other methods to eliminate conscious or unconscious bias on part of observer and subject; careful separation of symptomatic reactions from objective findings; psychologic appraisal of subject.

(*i*) *Duration of Observation.*—Treatment to continue until any intrinsically undesirable or harmful manifestations have had time to develop as well as until sufficient time has elapsed to demonstrate beneficial effects; comparison of rapidity of cure or improvement with that of other methods of treatment.

(*j*) *More Than One Clinical Investigation.*—Several different investigators working independently; conclusions to be made independent of one another's results.

2. During its 1961 hearings, the Subcommittee heard extensive testimony on the scientific basis for proving drug efficacy.[2]

Dr. Charles D. May, Professor of Pediatrics, New York University Medical School, testified to the "fallacy" of assuming that "a collection of impressions" will furnish the truth, pointing out that "this approach did not prevent doctors from having unbounded faith in the curative powers of leeches for hundreds of years before scientific evaluation became the preferred means of judging efficacy of therapy." He testified that the "magnitude of sales of a drug after vigorous promotion is no recommendation for its usefulness or efficacy" and recommended that "the

[2] Hearings on Drug Industry Antitrust Act before the Senate Subcommittee on Antitrust and Monopoly of the Senate Judiciary Committee, 87th Cong., 1st Sess., pursuant to S. Res. 52 on S. 1552, Pt. 1 (1961).

present concept of controlled methods of evaluation of drugs" be employed to determine efficacy.[3] He reasoned that "in modern times the methods required for scientific evaluation of a drug require design of experiments, controlled experiments, beyond the facilities of each and every individual physician,"[4] that "an orderly approach of experimental evaluation would be a vast improvement over wanton individual experience,"[5] and that efficacy is "unlikely" to be determined unless a drug is tested "by a systematic, scientific evaluation according to the best modern procedures."[6]

Dr. Louis Goodman, Professor of Pharmacology, University of Utah College of Medicine, and coauthor of "The Pharmacological Basis of Therapeutics," said that those who had seen "the mass of laboratory and clinical information submitted to the FDA, even by the very best drug houses, in support of new drug applications are repeatedly dismayed by the welter of anecdotal case reports and uncontrolled clinical studies;"[7] that the "average practicing physician—and I have helped to train hundreds of them—just does not have the time, the facilities, the skill, nor the training to be an expert in the determination of drug efficacy."[8] Dr. Goodman insisted upon "basic, original, clinical evidence that [a] drug is a useful drug and that the claims made by the manufacturer are valid." He noted that 35 out of 100 people react to placebos, and "that is why experts talk in detail about controlled clinical

---

[3] *Id.*, pp. 195–196.
[4] *Id.*, p. 204.
[5] *Id.*, p. 205.
[6] *Id.*, p. 207.
[7] *Id.*, p. 215.
[8] *Id.*, p. 217.

experiments; that is why the individual practicing doctor cannot be the judge." [9]

Dr. Louis Lasagna, head of the Division of Clinical Pharmacology at Johns Hopkins, testified that in "some areas of disease, it is easier to devise adequate tests than in others. Yet it remains a basic principle that in the absence of measurable evidence that two drugs differ, one must assume that the two drugs are alike. * * *

"Because I have often testified to the importance of double-blind controlled trials in clinical research, some may consider my statements to constitute a demand for uch trials on all new drugs. Not at all. There are certain areas of disease—insomnia, anxiety, depression, and pain, for example—where the absence of such controlled trials leads to, in my opinion, chaos. In some other areas—acute leukemia, for example—an answer might be easily obtainable in a small number of patients without the use of double-blind controls. The emphasis should be on scientifically acceptable evidence." [10]

On this point he asked: "What of the allegations that only by extensive clinical experience can the true efficacy of a drug be established? The history of medicine is, unhappily, replete with examples of useless drugs employed for years, decades, or centuries, by countless physicians before a few properly conducted experiments proved the drugs to be without value." [11]

Dr. Allan M. Butler, Professor of Pediatrics, Harvard Medical School, commented that "It is very hard * * * to have a private, practicing physician in his private practice properly appraise the efficacy and toxicity of a drug.

---

[9] *Id.*, p. 243.

[10] *Id.*, pp. 282–283.

[11] *Id.*, p. 283.

"It must be done by physicians who have the facilities to set up controlled clinical testing. * * * "[12]

Dr. Harry F. Dowling, Professor and Head of the Department of Medicine, University of Illinois, testified on minimizing patient and investigator bias, on placebo effect, and on double-blinding. He related the following example:

> It is a well-known fact that if the physician gives a patient a drug and tells the patient or otherwise implies to him that this drug is going to make him better, an appreciable percentage of the patients will indeed feel better and will insist that the drug has caused the improvement. This is the placebo effect.
>
> Thus, if a new tranquilizer drug were to be tested, a certain number of patients should be given an inert substance which looked exactly like the tranquilizer, and neither they nor the doctors evaluating the drug should know who received the dummy and who received the real drug. Only by this method could the efficacy of such a drug be tested properly.[13]

He emphasized that comparative analysis between patient groups is necessary "in order to determine that the differences in the case fatality rates were due to the [drug] and not due to chance alone"; that the "proposed bill would make sure that the proper tests had been done * * *."[14] He concluded:

> Anything that would improve the practice of medicine would benefit the patient. If more adequate tests are done to determine the efficacy of a drug, the patient benefits; if claims for a drug coincide with proved accomplishment, the patient benefits, and if the doctor knows what drug he is prescribing, what it will

[12] *Id.*, p. 355.
[13] *Id.*, p. 411.
[14] *Id.*, p. 412.

do and what the possible toxic reactions are, the patient obviously benefits.[15]

Dr. Maxwell Finland, Associate Professor, Harvard Medical School, observed that, with "respect to the use and evaluation of drugs, the practicing physicians must, as in other areas, rely on those with more experience and those who have special knowledge, training and experience and who have the facilities and opportunities to make the necessary observations and comparisons that are required to make a reliable judgment. If he is honest, the individual physician will admit that he can only gather impressions concerning the effect of the remedy he uses in an individual patient, but he cannot even in that case determine the likelihood that the response was a matter of chance or was due to any of the different methods of treatment that he employed, nor can he say what might have been the result were some other drug used. Only when adequate numbers of properly studied cases are gathered and the responses analyzed and compared with similar data from other well documented experiences can one gain some approximation of the true effectiveness of any remedy." [16]

Dr. Finland illustrated the need for adequate and well-controlled studies by giving the example of 100 reported staphylococcal pneumonia patients treated with drug X, none of whom died, although the best alternative drug available still resulted in 50% mortality. Drug X was hailed as a wonderful new treatment. The physician had done lab tests and determined the presence of staphylococcus before administering treatment, but no X-rays were taken to confirm the diag-

[15] *Id.*, p. 424.
[16] *Id.*, p. 430.

nosis. Upon closer scrutiny, no patient was found actually to have had staphylococcal pneumonia:[17]

> Now, this is the type of evidence that is presented to the FDA and which the FDA is told: "Now, look, you are taking thousands and thousands of cases and you are pitting these against a few studies." They do not say "careful" or "controlled" studies. They say "a few cases" by some experts, so-called, only people who think that they are experts. This is the sort of material they are presenting to you.
>
> This is the sort of thing that I say is dangerous because another doctor who knows how to make a diagnosis of staphylococcal pneumonia will use that drug to the peril of this patient.

Eugene N. Beesley, President of Eli Lilly & Co. and Chairman of the Pharmaceutical Manufacturers Association, testified that "we support the principle that, before introducing a new drug, the manufacturer should be required to submit to the Food and Drug Administration substantial evidence not only that the drug is safe but also that it produces the results claimed. * * * The test of effectiveness is to be whether the drug will produce the specific effects asserted by the manufacturer."[18] He stated that "[t]he key words are 'well-conducted clinical trials by competent clinicians.' "[19]

3. After enactment of the 1962 Amendments, Congress continued to reflect its interest in the requirement of adequate scientific evidence to show drug effectiveness. In S. Rep. No. 1153, 89th Cong., 2d Sess. (1966), the Committee investigating Interagency Drug Coordination made the following observations (pp. 197-200; footnotes other than 86a omitted):

---

[17] *Id.*, pp. 450–451.

[18] *Id.*, Pt. 4, pp. 1996–1997. See also pp. 1998–1999.

[19] *Id.*, p. 2007. See also p. 2012.

5. *Authorities' Insistence on Controlled Clinical Trials.*—Scientific authorities believe that, with certain exceptions, efficacy of a drug should be proven on the basis of a controlled clinical experiment; i.e., an experiment based upon comparison of results—those obtained by the drug as contrasted with those obtained, in a comparable group of patients, by either a blank pill (with inert ingredients; i.e., a placebo) and/or standard therapy.[86a]

University authorities on testing hold to the view that a drug may be regarded as truly *efficacious* only if its use yields better results (in a significant statistical sense) than those obtained from a placebo in such an experiment.

* * * * *

A careful scientific experiment requires, they note, other elements; e.g., (*a*) careful biostatistical preparation and analysis; (*b*) *randomized selection* of similar patients into groups which are to receive the drug being tested and the placebo; (*c*) careful withholding of knowledge from the patient as to whether he is being given an active drug or placebo; (*d*) careful efforts to withhold from the investigator the identity of the drug (if the experiment is to be, by definition, of a double-blind character).

Danger of Bias.—These elaborate procedures have been devised as a means of overcoming a wide variety of factors which, over the years, had led to erroneous belief in therapy which later proved to be worthless. Most notable among these factors has been what is known as the placebo effect. This refers to the potentially powerful

[86a] Comparison is not always either necessary or possible. For example, a physician treating a patient with a disease whose course is certain to be fatal and for which no alternative therapy is available, might decide that it would be contrary to medical ethics to give a mere placebo to a "control" group of such doomed patients. Most clinical conditions do not, however, involve circumstances ruling out a control.

effect of belief in a drug. It derives from suggestibility of a patient—his expectation that a drug will "cure" him. * * *

A second factor which the controlled experiment is designed to eliminate is the bias of the investigator, intentional or unintentional.

Most such bias is unwilling on the investigator's part. It derives from his own desire to discover a genuinely efficacious drug. His enthusiasm tends to "color" his observations.

* * * * *

6. *Dispute Over Science by "Majority Vote."*—Many university authorities contend that credence cannot be placed in what are termed "clinical impressions" no matter how many hundreds or even thousands of physicians may offer them as "proof" of a drug's value.

* * * * *

A dozen years ago, a distinguished British scientist, Sir Austin Bradford Hill, stated: * * *

> Two or three uncontrolled observations may give, merely through the customary play of chance, a favorable picture in the hands of one doctor, an unfavorable picture in the hands of a second. And so the medical journals, euphemistically called the "literature," are cluttered up with conflicting claims—each in itself perfectly true of what the doctor saw, and each insufficient to bear the weight of the generalization placed upon it.

Professor Louis Lasagna has criticized supporters of such "clinical experience" * * *:

> There is a common belief that the addition of many bits of disorganized data will somehow yield a clear picture, that the multiplication of zeros (or near zeros) will somehow produce an important finite good. Unfortunately, the his-

tory of therapeutics gives scant support to such philosophy. In medicine, theories and therapeutic practices, including those espoused by the majority come and go. One generation bleeds the ill, another scoffs at bloodletting. One generation insists on prolonged bedrest, another preaches the dangers of immobilization and the benefits of early ambulation for everything from surgery to cardiac infarct. * * * There thus can be no sense of confidence automatically generated by "traditional" practice; in therapeutics, as in many other areas of human endeavor, there is no magical safety in numbers.

It is a sad fact that the individual physician cannot really gauge the true worth of many remedies. There are, to be sure, clinical situations where the course of the illness is so stereotyped and predictable that a drug-induced change is readily appreciated. In such cases, the doctor can be his own judge and jury. But in many situations, the patient's disease has spontaneous and erratic ups-and-downs or is self-cured at unpredictable rates of improvement or responds dramatically to suggestion and the impact of the doctor-figure. In these instances, it is difficult enough to assess remedies with a strict experimental design; it is sheer bravado to draw conclusions without such a plan.

On another occasion, Dr. Lasagna stated * * *:

* * * it is almost certain that the "experts" do not subscribe to the AMA's democratic notion that "only the medical profession, after widespread usage, can determine the true effectiveness of a drug." *Scientific truth is not arrived at by majority vote. There is too much evidence that the medical profession, like*

499-736—73——4

the rest of us, can be misled by "practical experience" to encourage the testing of truth by referendum. * *

The amassing of unsubstantiated views of practitioners is regarded as inconclusive and generally unworthy of scientific debate. Critics liken such unworthy "evidence" as equivalent to "testimonials," such as medicine, itself, long fought in patent medicine advertisement.

Debate Over Antibiotic Combinations.—University experts criticize, for example, what they regarded as unjustified medical protests which poured in on FDA because an expert advisory committee objected to antibiotic combinations for treatment of upper respiratory infections. The experts assert that no matter what the number of protests, 400, 1,000, or more, such subjective views by M.D.'s cannot possibly refute the advisory committee's judgment that there is no scientific proof of these combinations' efficacy. University experts feel that FDA should not be so easily influenced by masses of protests, considering that the protests may not contain a shred of scientifically controlled evidence.

APPENDIX B

The Scientific Principles Underlying Adequate and Well-Controlled Clinical Investigations To Prove Drug Effectiveness

The Federal Food, Drug, and Cosmetic Act requires that all drugs be effective for their labeled indications. A drug must either be generally recognized as effective or approved for effectiveness by FDA.[1] To explain the basis upon which FDA determines effectiveness, the Commissioner of Food and Drugs published regulations outlining principles developed over a period of many years by investigators actively engaged in the evaluation of the therapeutic efficacy of drugs.[2] This appendix summarizes the scientific basis for these principles.

Although in the earlier part of this century there were occasional calls for more scientifically acceptable methods in the clinical evaluation of drugs,[3] development of the principles of the controlled clinical trial and their actual application did not gain impetus until the time of World War II.[4] Great strides

[1] 21 U.S.C. 321(p), 355 (d) and (e).

[2] 21 C.F.R. 130.12(a)(5).

[3] Irons, *The Clinical Evaluation of Drugs*, 93 J. Am. Med. Assn. 1523–1524 (November 16, 1929); Leake, *The Pharmacologic Evaluation of New Drugs*, 93 J. Am. Med. Assn. 1632–1634 (November 23, 1929).

[4] Atkins, *Conduct of a Controlled Clinical Trial*, Brit. Med. J. no. 5510, at 377–379 (August 13, 1966); Van Winkle et al., *Laboratory and Clinical Appraisal of New Drugs*, 126 J. Am. Med. Assn. 958–961 (December 9, 1944); Gold, "Experience in Human Pharmacology," in *Quantitative Methods in Human*

have been made in therapy in the past without the benefit of this experimental device, simply on the basis of the uncontrolled observations of astute clinicians. Modern controlled clinical trials were not needed to establish that such important drugs as digitalis, ether, penicillin, quinine, and the sulfonamides had therapeutic effectiveness. Controlled trials were, however, needed to define the exact therapeutic roles of these drugs. Further, thousands of drugs which, on the basis of "clinical experience," were once accorded an "indispensable place" in therapy, are now known to be useless as the result of controlled trials. Many of these subsequently discarded agents and therapies (gold therapy in tuberculosis, violent purging, bleeding, total colectomies for "autointoxication") lingered on for great periods of time in the practice of medicine, and were not only inefficacious, but seriously injurious or occasionally lethal for some patients.[5]

The function of the formal controlled clinical trial is to separate the relative handful of discoveries which prove to be true advances in therapy from a legion of false leads and unverifiable clinical impressions, and to delineate in a scientific way the extent of and the limitations which attend the effectiveness of drugs. The importance of properly conducted, controlled tests as the only reliable method of evaluating drug effectiveness in most instances is a theme constantly expressed in modern writing on pharmacology:

> Let me try to recapitulate briefly. Why are controlled trials important? Because few dis-

---

Pharmacology and Therapeutics (1959); Gaddum, *Clinical Pharmacology*, 47 Proc. Royal Soc. Med. 195–204 (1954); Barber, *Drugs and Society* 19 (1967).

[5] Atkins, *supra;* Pickering. *The Place of the Experimental Method in Medicine*, 42 Proc. Royal Soc. Med. 229–234 (1949).

eases run a course which is precisely predictable; because patients, doctors, and details of medical care differ greatly from place to place, and from one time to another; and because most patients and most physicians are biased towards expecting therapeutic benefit. A properly designed trial is, therefore, an attempt to safeguard the investigator from unwarranted conclusions. The double-blind aspect of experimental design is only one facet of the problem. The principles of concurrent comparison of treatments, of unbiased allocation of patients to treatment groups, and of statistical analysis of data are equally important. Will adherence to these principles insure a successful experiment? Of course not. One still has to ask an appropriate and answerable question. But the chances of reaching a correct answer are considerably greater when the above principles are adhered to.[6]

* * * * *

In order to decide whether patients treated in one way are benefited more than those treated in another, there is no possibility of avoiding the use of numbers. The mere statement by a clinician that patients do better with this or that treatment is due to his having formed an opinion that more patients are helped by the treatment he advocates than by other treatments. The opinion is based on numbers, but having omitted to record exactly how many patients have been treated by different methods and having omitted to ensure that the only variable factor affecting the patient was the treatment in question, only a "clinical impression", instead of a scientific fact, can be stated. This is a pity, for progress is delayed when convinced opinions are offered in place of

[6] Lasagna, *Controlled Trials: Nuisance or Necessity*, 1 Meth. Inform. Med. 79, 81 (1962).

convincing facts. The former, though not necessarily wrong, are unreliable, despite the great assurance with which they are often advanced.[7]

* * * * *

It is the judgment of competent authorities that the relative effectiveness of specific therapeutic measures can be reliably established only in well-controlled comparative clinical investigations. These embody the following essential procedural steps in the design of the study: descriptions of all the patients in the total sample; random allocation of them so as to provide adequate, suitable and comparable control as well as therapeutic groups; standardized criteria of therapeutic effectiveness.[8]

* * * * *

The history of medicine abounds with remedies that were long and widely used before falling into disrepute and vanishing. A designed test might have greatly hastened their fall from favor and have thereby encouraged the search for something better.

It is these situations that the controlled clinical trial is designed to meet.[9]

* * * * *

Well controlled clinical evaluations reflecting the above ideas are not the invariable rule in practice today. Indeed, some resistance to the very concept often arises. The traditional authoritative attitudes held by most physicians make it difficult for them to believe that they cannot put complete trust in their individual

---

[7] Laurence, *Clinical Pharmacology* 11 (2d. ed. 1963).

[8] Committee on Public Health, New York Academy of Medicine, *The Importance of Clinical Testing in Determining the Efficacy and Safety of Drugs*, 38 Bull. N.Y. Acad. Med. 415, 426 (1962).

[9] Hill, "Aims and Ethics," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 3 (1960).

judgments but must defer to the conclusions of some nearly unknown investigator. However, the individual physician combining the responsibilities of evaluation with those of patient care is easily misled by "mere experience," and an acceptable evaluation can be done only by executing a carefully prepared experimental design.[10]

* * * * *

What constitutes an adequately controlled clinical trial necessarily varies, depending upon the drug effect being evaluated. The more important general requirements for all trials are an appropriate and adequately sensitive method of evaluation, an adequate number of subjects, lack of bias, concurrent comparison of the new drug with a reference drug over a range of doses, and appropriate statistical validation. Depending upon the drug effect being evaluated, sore clinical trials must be conducted under so-called blind conditions.[11]

* * * * *

We have what we call our clinical experience. The difficulty about clinical experience as I have known it in my own case and vicariously in the case of my colleagues is that, in a general way, it is unplanned and haphazard. We are victims of the freaks of chance. We may have a little run of experience which convinces us that a certain treatment is the most useful. We are also victims of the attractive propaganda leaflets that arrived from the drug companies the week before and of the persuasive eloquence of our colleagues. And so, working in this way, we often, in good faith, carry out a treatment for a considerable time and at the end of it we do not

[10] Meyers et al., *Review of Medical Pharmacology* 15 (1968).

[11] Fingl & Woodbury, "General Principles," in Goodman & Gilman, *The Pharmacological Basis of Therapeutics* 1, 30 (1965).

know whether our treatment has done our patients any good or not * * *

To my mind, then, the clinical trial is the best way of getting to know what are the probabilities that a given form of treatment will affect one's patient for better or for worse.[12]

* * * * *

The field of medical observation, it is necessary to remember, is often narrow in the sense that no one doctor will treat many cases in a short space of time; it is wide in the sense that a great many doctors may each treat a few cases. Thus, with a somewhat ready assumption of cause and effect and, equally, a neglect of the laws of chance, the literature becomes filled with conflicting cries and claims, assertions and counterassertions. It is thus, for want of an adequately controlled test, that various forms of treatment have, in the past, become unjustifiably, even sometimes harmfully, established in everyday medical practice.[13]

In the discussion that follows, we outline the ten scientific principles that experts in pharmacology have enunciated as the requirements for and characteristics of an adequate and well-controlled clinical investigation. These principles form the foundation of and are embodied in the Commissioner's regulations interpreting the drug effectiveness evaluation standards. See also, Doll, "Clinical Trials," in Zaimis & Elis, *Evaluation of Drugs in Man* 159 (1965); Sheps, *Problems in Clinical Evalation of Drug Therapy*, 5 Perspectives in Biol. Med. 308, 361 (1962).

1. *There must be a plan or protocol for the study.*

---

[12] Pickering, in Council for International Organizations of Medical Sciences, *supra*, at 164.

[13] Hill, *Statistical Methods in Clinical and Preventive Medicine* 4 (1962).

It is impossible to make any meaningful evaluation of the evidence from a clinical trial unless the reviewer has a detailed plan or protocol of the study at his disposal. This protocol must explain the objectives of the study, the complete methodologic details of how the study was done, and how the data were collected and analyzed.

> As the controlled clinical trial stands or falls on its plan, the report must describe this in detail. It is important to give the diagnostic criteria employed, so that it is evident that the patients selected were homogeneous in respect of their illness, and so that it is obvious in what clinical conditions the results of the trial might later be applied. The recruitment of patients, whether for example from all admissions to the wards or out-patients which satisfy the diagnostic criteria, or from some special source, and the method of allocation to Treated and Control groups, require special attention on the report as this is the crucial stage in a trial. The therapeutic regime to be given to both groups must be stated, in particular whether the dose is to be fixed or left to the physician's discretion, together with the reasons for this decision * * * The plan also includes the choice of clinical and laboratory tests used to measure the patient's progress, their timing and the scales by which they are to be interpreted, and the arrangements made to ensure equality of observation, both objective and subjective, on the Treated and Control groups.[14]
>
> * * * * *
>
> . . . a written plan should be made which states the objects of the trial, the principles underlying the plan, the criteria for admitting

[14] Knowelden, "The Analysis and Presentation of Results," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials--A Symposium* 155 (1960).

patients, exact details of how to proceed and instructions about what information is to be collected and how it is to be recorded. Specimen documents for the recording of the data should be attached to the plan. . . .[15]

(2) ***The plan or protocol must include a clear statement of the objective of the study.***

The importance of a clear formulation of the basic question or questions to be answered by the clinical trial is emphasized by many authorities.

> The first step in the controlled trial is to decide precisely what it sets out to prove. It is essential that initially its aims should be laid down in every detail.[16]
>
> * * * * *
>
> The precise goals and objectives of any clinical research project must be clearly specified at the very outset * * * Only by formulating the correct questions can the monitor and the investigator achieve the answers they seek. And it is these answers that are the primary objects of the clinical trial.[17]
>
> * * * * *
>
> The choice of the primary objects of the trial is of profound importance. They should be chosen after considering what is already known of the course of the disease and the effects of established treatments.[18]
>
> * * * * *

---

[15] Truelove, "Therapeutic Trials," in Witts, *Medical Surveys and Clinical Trials* 148, 150 (2d ed. 1964). See also Cox, *Planning Clinical Experiments* 55–72 (1968); Hill, *Statistical Methods in Clinical and Preventive Medicine* 25 (1962).

[16] Hill, *Principles of Medical Statistics* 251 (8th ed. 1966).

[17] Taber, *Proving New Drugs* 78 (1969).

[18] Truelove, *supra*, at 149.

The first step in the design of a therapeutic trial is the formulation of the questions that it is hoped to answer. It is wise to limit the number of questions and to make these few absolutely precise. This has the disadvantage that the answers are limited to specific questions and are unsuitable for generalizations. On the other hand, if the questions are made too complex, the investigator will be unable to draw any firm conclusions at the end of the trial. He will be faced with a number of inconclusive answers, each based upon too few observations. In short, the exact aim of the trial should be thought out in detail before it is begun. This will involve such points as the accurate description of the patients to be included in the trial, the treatment(s) that they are to be given, and the measurements that are to be made to reveal the progression of their illness and the effects, if any, of the drug.[19]

(3) *The plan or protocol must include a method of selection of the subjects that provides for adequate confirmation of the disease state present, including criteria of diagnosis and appropriate confirmatory laboratory tests.*

Since the clinical trial is intended to predict on the basis of experience with a small sample of patients with a given disease, what the response of the generality of patients with that disease will be to a specific treatment, it is essential that the protocol define the disease with extreme care. Accurate diagnosis

[19] World Health Organization, *Principles for the Clinical Evaluation of Drugs*, Tech. Rep. Ser. No. 403, at 21 (1968). See also Laurence, *supra*, at 15; Lasagna & Meier, "Experimental Design and Statistical Problems," in Waife & Shapiro, *The Clinical Evaluation of New Drugs* 37, 38 (1959); Hill, *The Clinical Trial*, 190 Practitioner 85, 86 (January 1963); Cox, *Planning Clinical Experiments* 14, 57 (1968).

of the disease is also important. Since the diagnosi of most disease entities is not absolutely clear cu the particular diagnostic criteria used by the investi gator must be presented. In those diseases in whic laboratory tests may prove of material aid in accurat diagnosis, these tests should be done and their result reported.

> The first area to work upon will be the exac definition of the clinical problem to be tackled The patient and disease expected to be deal with must be identified accurately for the re sults to be meaningful to others. It is not enough to talk vaguely of hypertension or wound infection or rheumatoid arthritis, for there are multiple degrees and stages of disease. Without precision of original definitions there can be no precision transmitted in the conclu sions, no matter how carefully all the interme diate steps have been measured.
>
> The definition of the disease must be written down, and the necessary and sufficient criteria for a diagnosis accepted. Clinical diagnosis are not always possible in clear-cut terms, for con flicting evidence must be weighed, and absent features considered in relation to the grossness of those features present. Yet, if this experi ment is to be meaningful to others who did not examine these patients, then some definite de siderata must be met in all patients.[20]
>
> * * * * *
>
> The problem of diagnosis and exclusions is broad and recurrent. Diagnosis is a difficult art, and the problem of correct diagnosis is worse, not better, when one is conducting a thera peutic trial. If the patient's response to therapy is used to judge the usefulness of that therapy in future patients who have the same disorder as the one incorrectly diagnosed in the first

[20] Cox, *supra*, at 16.

patient, a great deal of future mischief may result. One must therefore set up criteria for diagnosis and exclusions that will prevent admission of patients with diseases that are superficially similar to, but fundamentally different from, the disease under study.[21]

* * * * *

Regardless of the size of the aggregation, therapeutic comparison cannot be fruitful unless the two groups of patients to be studied are more or less homogeneous. The chief prerequisites for homogeneity are uniform etiology and pathogenesis in the two comparison groups.[22]

* * * * *

. . . the investigator will need to consider such questions as accuracy of diagnosis and severity of disease in patients to be admitted to the trial, whether the patients have a history of therapy that might medify the course of their illness and thus, possibly, confuse the issue of the trial, whether they should be limited to defined area (e.g., excluding the very young or old), whether they should be without other diseases than the one under investigation, and so on.

The criteria must not, of course, be made so rigid as to limit unduly the patients available for the trial, nor to make the answer to the trial so narrow as to be of little use in medical practice.[23]

---

[21] Carr, *Practical Considerations in the Design of Clinical Trials*, 98 Canad. Med. Assn. J. 307, 311 (February 10, 1968).

[22] Martini, "Experimental Design in Clinical Medicine," in Ladimer & Newman, *Clinical Investigation in Medicine: Legal, Ethical and Moral Aspects* 398, 400 (1963).

[23] World Health Organizations, *supra*, at 21. See also Taber, *supra*, at 80. Truelove, *supra*, at 150, Fletcher, "Criteria for Diagnosis and Assessment in Clinical Trials," in Council for International Organizations of Medical Sciences. *Controlled Clinical Trials—A Symposium* 19, 20, 27 (1960).

The considerations relevant to the definition of the population of patients used in a study have been discussed,[24] and the problems of the taxonomy of disease and how these relate to clinical trials have been explored, in great depth.[25]

(4) *The plan or protocol must include a method of selection of subjects that provides for assignment of the patients to test groups without bias.*

The concept that patients should be allocated to treatment groups in such a way as to minimize the risk of bias is central to the controlled clinical trial. In the words of one researcher, "It is difficult to conceive of a rule more important for the clinical investigator than that of avoiding bias in the allocation of cases to different treatment groups." [26] In the framework of the clinical trial, "bias" is introduced into the experiment to the extent that known or unknown variables capable of influencing treatment outcome are unequally distributed between treatment groups. Bias is thus something more than the conscious or unconscious prejudices of the investigator, observer, or subjects.[27]

The distinctive feature of an experiment is that the investigator comparing the effects of two or more factors assigns them himself to the individuals who comprise the test population, and he can assign these factors in such a way as to reduce the risk of bias to a predetermined degree.[28] Reducing the risk of bias in-

---

[24] Mainland, *Elementary Medical Statistics* 29–38 (2d ed. 1963).

[25] Feinstein, *Clinical Judgment* (1967).

[26] Lasagna, *The Controlled Clinical Trial: Theory and Practice*, 1 J. Chronic Dis. 353, 356 (1955).

[27] Houde et al., "Clinical Measurement of Pain," in deStevens, *Analgesics* 75, 80 (1965).

[28] Mainland, *Elementary Medical Statistics* 19 (2d ed. 1963).

volves more than the mere matching of attributes known to be capable of influencing results, for this would not ensure that hidden variables are equally distributed. Only deliberate randomization ensures the chance distribution of hidden variables. If randomization is properly performed, and if chance can be ruled out as the likely cause of observed treatment difference, then the cause of the difference can reasonably be attributed to the treatment. On the other hand, if proper randomization in the allocation of patients to treatment groups has not been performed, an observed difference in the response of the treatment groups, even though statistically "significant," may be due either to the treatments or to the bias. In such a circumstance, one cannot reasonably attribute the observed difference solely to treatment differences, and hence a conclusion of treatment efficacy is unwarranted.

There are a few circumstances in which the random allocation of patients to treatment groups is not absolutely necessary, but these almost always correspond to those situations in which a historical control would be considered a satisfactory standard of comparison for a test treatment. No expert in the fields of biostatistics or clinical pharmacology disputes the crucial importance of the method of assignment of patients to treatment groups in determining whether a particular clinical trial constitutes an adequate and well-controlled investigation.

> Method of allocating patients to different groups, or of assigning different treatment to different patients, are subject to both conscious and unconscious bias if they depend on a physician's choice. Some method of random allocation is essential to avoid this danger. Commonly, this is achieved by the use of tables of random numbers, where each digit or each combination

of a given number of digits has an equal probability of selection * * *

The purpose of randomization is twofold: (1) to render the subject groups as equivalent as possible in regard to all variables other than the treatment under study, and (2) to provide scientific justification for the application and interpretation of statistical tests of probability and significance.[29]

* * * * *

One noticeable feature of the modern controlled trial is the allocation of patients to treated or control group by some entirely random method. . . .

The object is to produce two (or more) groups of patients similar in respect of any known or unknown characteristics that may affect the prognosis of their illness but differing in their treatments.[30]

* * * * *

Randomized designs, in which the order of entry of patients into a study is determined by chance, are often more useful than systematic designs, with preset order of entrance into the trial, for systematic designs are liable to systematic errors; i.e., the selection system may have a built-in bias, unknown to the investigator.[31]

* * * * *

The best way of getting equivalent groups is by allocating patients to them by "random allocation". To allot patients alternately or otherwise systematically is not satisfactory as the physician almost inevitably knows into what treatment group a patient will go whilst engaged in deciding whether the patient should enter the trial, and he may be unconsciously

---

[29] World Health Organization, *supra*, at 22.

[30] Hill, *The Clinical Trial*, 190 Practitioner 85, 88 (January 1963).

[31] Carr, *supra*, at 309.

influenced by this if he has strong feelings about either the patient or the value of the respective treatments. With random allocation the treatment group into which the patient goes is only discovered after it as been decided to enter him in the trial.[32]

* * * * *

The patients are assigned treatments randomly—from a table of random numbers, or by birth date (or month)—odd dates on treatment A, even dates treatment B, or by a roulette wheel. There is no valid reason for the investigator to know what the randomization scheme is. Patients are not assigned alternately. They are not assigned haphazardly.[33]

* * * * *

No high-powered analysis of data will compensate for data of poor quality or dubious validity. A prime assumption in the statistical analysis of the comparative performance of two drugs is that there has been no bias in the allocation of patients to one particular treatment group. It is impossible to interpret a "statistically significant difference" between drugs if the therapeutic challenge has been made more difficult for one treatment by purposeful shunting of even occasional patients into that group. In such a case, Drug A may look better than Drug B either because it is better than B, or because the patients getting Drug A were less ill than those getting B. Since there is no way to distinguish between these alternatives, such an experiment is usually worse than useless.

The "random" assignment of patients required for a valid comparison must not be interpreted in the lay sense of "haphazard."

---

[32] Laurence, *supra*, at 15.

[33] Schneiderman, "Defensive Design of Clinical Trials," in Sternberg & Newcomer, *The Evaluation of Therapeutic Agents and Cosmetics* 66, 68, (1964).

> Schemes such as alternation of subjects or odd and even serial numbers are notoriously subject to bias in assignment of patients. That they are often a minor gain in convenience is poor payment for the loss of an objective measure of error. Although an investigator may convince himself that there is no conceivable way in which the serial number could affect the characteristic under study, this state of affairs may merely be a tribute to his lack of imagination. It cannot be emphasized too strongly that the rationale for confidence statements (vide infra) is completely dependent on the employment of true randomization.[34]
>
> * * * * *
>
> "This is achieved by some sort of randomizing process for assigning patients to the two groups * * *
>
> It is worth emphasizing that it is better to build up a well-controlled series in this way even if it has to be done gradually than to get a larger series poorly controlled. Large numbers in themselves are worse than useless if the groups are not comparable, since they encourage confidence in an erroneous opinion.[35]

* * * * *

Other scientists who have considered this problem emphasize these same principles.[36]

---

[34] Lasagna & Meier, "Experimental Design and Statistical Problems," in Waife & Shapiro, *The Clinical Evaluation of New Drugs* 37, 44 (1959).

[35] Marshall & Merrell, *Clinical Therapeutic Trial of a New Drug*, 85 Bull. Hopkins Hosp. 221, 224 (1949).

[36] Martini, *supra*, at 401; Deeley, *The Random Allocation of Patients in Clinical Trials*, 5 Meth. Inform. Med. 100 (April 1966); Armitage, "The Construction of Comparable Groups," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 14 (1960); Truelove, *supra*, at 151; Taber, *supra*, at 89; Goldstein, *Biostatistics: An Introductory Text* 6 (1964); Goldstein, et al., *Principles of*

(5) The plan or protocol must include an outline of the methods of quantitation and observation of the parameters studied in the subjects.

This principle simply requires a statement of the index or indices of drug effect, or the "index of therapeutic accomplishment," used in the study, and a description of how these indices were measured. The selection of a suitable index of effect is a most important part of protocol planning and may be either simple or relatively complicated, depending on the disease entity or symptom being treated.

> The choice of variables to be measured will depend upon the disease. In some, for instance, the status of the heart will be of dominating importance, in others joint swelling and pain, and so on. The methods of measurement must remain unchanged throughout the course of the trial. Any criteria of assessment of the patient's condition must similarly remain unaltered. Clear definitions of methods, criteria, times of measurement, and clinical assessment must be agreed upon before the trial starts. Such methods, criteria and time schedules should be adhered to as closely as possible. Every departure from the rules lowers the efficiency of the trial.
>
> In the field of psychopharmacology, where the variables involved make objective measurements either difficult or irrelevant, a scoring system for the evaluation of the symptoms may prove of value. The use of newer techniques and instruments may in time aid quantitative evaluation of psychiatric signs and symptoms.

Drug Action 799 (1968); Nickerson, *The Control of Unknown Variables*, 97 Canad. Med. Assn J. 118–122 (July 15, 1967); Mainland, *Elementary Medical Statistics* ch. II & III (2d ed. 1963); Hill, *The Clinical Trial*, 247 N. Eng. J. Med. 113, 115 (1952); Hill, *Statistical Methods in Clinical and Preventive Medicine* 9, 20 (1962); Doll, *supra*, at 161; Gaddum, *supra*,

It is axiomatic in any trial that the same care and precision in measurement be applied to all groups whether treated with a new drug or not.[37]

* * * * *

Similarly, we must plan in advance the observations that must be made to measure the changes in the patient that follow the exhibition of the treatment. We must closely consider the advantages and limitations of such objective but isolated measurements as the changes in the sedimentation rate and temperature, and the advantages and limitations of the more subjective clinical assessments of the changes in the "whole" patient. The stress laid upon such observations being made "blind", i.e. without a knowledge of the treatment given to the individual patient, is merely to ensure that the observer can make his assessments without fear of bias, conscious or unconscious, and, equally, without fear of being accused of bias.[38]

* * * * *

When objective measurement of effects is possible, and when pharmacodynamic actions are potent, reproducible, and are not significantly influenced by psychic forces, precise evaluation is relatively simple. Drug actions which must be evaluated in terms of subjective responses, and especially those which are not in themselves impressive, are far more difficult to measure or even estimate. Drug actions on function, which, like blood pressure, have a tendency to wide spontaneous variation or are altered by immediate circumstances, such as tension, position, strain, etc., provide serious problems in assessment.[39]

---

[37] World Health Organization, *supra*, at 21.

[38] Hill. *The Clinical Trial*, 190 Practitioner 85, 87 (January 1963).

[39] Modell, *The Sentivity and Validity of Drug Evaluations*. 1 Clin. Pharmacol. Ther. 769, 770 (1960).

* * * * *

Once the patients are properly allocated to treatment groups, attention must be turned to errors of meaurement. An experiment excellent in all other respects may fail because measurements are made with insufficient care or because of inadequate record-keeping.

If measurements are not carefully made, real differences between drugs may be missed. Randomization and double-blind techniques will prevent the experimenter from frequently claiming differences that are not present. However, they will not help him to control poor precision of measurement. Unbiased but sloppy measurements may lead to the correct assertion that "no significant difference was found."[40]

* * * * *

* * * [a] well-designed trial may fail to provide a clear result because the techniques used for assessment are so prone to error that by their use the effect of the treatment, though real, cannot be detected. If confusion is to be avoided it is necessary for those who undertake clinical trials to be as careful over the choice, description, and execution of the methods they use for diagnosis and assessment as they are over other aspects of the design of their trial.[41]

The proper selection of indices of therapeutic response and the measurement of these indices is of such importance in the clinical trial that whole chapters of books have been devoted to this issue.[42]

---

[40] Lasagna & Meier, "Experimental Design and Statistical Problems," in Waife & Shapiro, *The Clinical Evaluation of New Drugs* 37, 45 (1959).

[41] Fletcher, *supra*, at 19. See also Marshall & Merrell, *supra*, at 226; Hill, *The Clinical Trial*, 247 N. Eng. J. Med. 113, 117 (1952).

[42] Feinstein, *supra*, ch. 14: Cox, *supra*, ch. 11 & 12.

(6) *The plan or protocol must include a description of the steps taken to document comparability of variables, such as age, sex, duration of disease and use of drugs other than those being studied.*

Since, as was pointed out above, the object of the controlled clinical trial is to be able to draw a causal connection between any observed difference in response of the treatment groups and the treatments under test, it is desirable to have any other variable which is likely to influence treatment outcome distributed among treatment groups in a relatively equal way. With large treatment groups, simple randomization will usually ensure this, but as treatment groups become smaller, there is a greater likelihood that a variable relevant to treatment outcome may be unequally distributed among treatment groups simply due to the play of chance. To minimize this risk, it is often wise to stratify or pair subjects before randomization for those characteristics likely to influence treatment outcome. If, on examining the treatment groups after a trial, a significant difference in the distribution of some variable is noted, one must consider if this variable is of a nature likely to influence treatment outcome. If so, techniques such as subdivision of cases may be applied to gauge the impact of this variable on treatment outcome.[43]

> The intention in a controlled trial is to provide groups of patients, alike as far as possible, except that they receive different treatments. Before the progress of the groups can legitimately be compared it is important to see that they were initially alike. The analysis must therefore include a section on comparability of groups. On general grounds this should usually show equality in age- and sex-distribution.

[43] Cox, *supra*, at 223.

The frequency of presenting symptoms, duration of illness at the start of treatment, number of previous attacks, and severity, are characteristics which can often come into such comparisons."[44]

* * * * *

If relevant background information on different variables in patients is available, "stratification" of such information may increase specificity of comparisons and conclusions. In other words, the original sample of patients may be sub-divided (stratified) into appropriate and more homogeneous sub-groups, and a random sample withdrawn from each of these for allocation to treatment. The sub-sample may or may not be proportional to the number of units in the sub-group."[45]

* * * * *

The object is to produce two (or more) groups of patients similar in respect to any known or unknown characteristics that may affect the prognosis of their illness but differing in their treatments. When that prognosis may vary considerably between patients of known characteristics, e.g. men compared with women, young people compared with old, patients presenting with fever compared with the afebrile, then it is wise and customary to subdivide the patients on entry to the trial by these characteristics and to make the random allocation within each such subdivision. Thus equality in these characteristics is predetermined.

With the unknown characteristics that may influence prognosis we can, of course, only rely upon the random distribution of the patients to equalize the groups in those respects. This is a reasonable assumption when the numbers involved are not too small and is, indeed, the aim

[44] Knowelden, *supra*, at 157.

[45] World Health Organization, *supra*, at 22.

> of the method. Thus, having first analysed the data to show that in observable characteristics the groups are indeed similar upon entry to the trial, we may with confidence explore the changes that take place in them subsequently and attribute any notable differences to the one known difference between them: i.e. the treatment. For such comparisons between the groups the indices customarily and traditionally used in medicine will be required."[46]

* * * * *

(7) *The plan or protocol must include a description of the methods of recording and analysing the patient response variables studied.*

This requirement, a logical extension of the foregoing principle, stipulates that the methods of measuring patient response shall be specified. The data generated by these measurements must be recorded and analysed, and knowledge of how this was done is vital to any judgment on the significance of the investigator's conclusions. There seems to be no disagreement in the literature as to the importance of this principle, and a vast body of literature exists describing appropriate methods of processing and analysis for particular types of data.

* * * * *

> Before a therapeutic trial is begun, a written statement should be prepared setting out in detail the question(s) being asked, the treatment(s) to be used and the measurements that will be required. This statement should incorporate information about the patients to be admitted and how they will be allocated to the

---

[46] Hill, *The Clinical Trial*, 190 Practitioner 85, 88 (January 1963). See also Marshall & Merrell, *supra*, at 225; Hill, *Principles of Medical Statistics* 260 (8th ed. 1966); Armitage, "Statistical Aspects of Clinical Trials," in Zaimis & Elis, *Evaluation of Drugs in Man* 165, 169 (1965).

different treatments. The method of making measurements must be laid down and the times at which they will be made must be specified. In all this the medical statistician should be closely concerned. In some matters laboratory staff should also be consulted, since additional work may be created by tests for toxicity or drug analyses.

Following this statement of the whole project, it will be necessary to construct forms upon which all the data of the trial will be recorded. Care here is of the utmost importance. Attention must be given to such matters as patient identification, the treatment administered to each, and the measurements and clinical assessments that will be made at defined intervals.

Skilful design of this form will save much time both in the recording and in the subsequent analysis of the data. It may be necessary to construct the form so that the data can easily be transferred to punch-cards and sorting machines or to a computer.[47]

* * * * *

In the adoption of other measures of the effect of treatment, detailed planning must as usual play its part. Before the trial is set under way it must be laid down, for example, precisely when and how temperatures will be taken, when full clinical examinations will be made, and what will be specifically recorded, how often and at what intervals X-rays will be taken. Standard record forms must be drawn up, and uniformity in completing them stressed. Unless these rules and regulations are well kept and observed by the clinicians in charge of the trial, many and serious difficulties arise in the final analysis of the results, * * *[48]

* * * * *

[47] World Health Organization, *supra*, at 27.

[48] Hill, *Principles of Medical Statistics*, 259 (8th ed. 1966).

499-736—73——6

> The design of patient report forms is largely predicated upon a proper experimental plan. They serve as the backbone of documentation: They are essential for the proper organization, analysis, and interpretation of data, as well as for the collation, storage, and retrieval of information from clinical studies. In our experience, one of the best means to satisfy good experimental design is to devise and prepare a patient report form with practicality and foresight, preferably designing it when the clinical protocol is devised—both become interlocked."

(8) The plan or protocol must include a description of the means of excluding minimizing bias from the observations.

Not only is it important in a controlled trial to avoid bias in the course of setting up the treatment groups to be compared, but it is also important to minimize the introduction of bias which may arise during the course of the experiment due to the preconceptions and expectations of the investigator, observer, and patient. This is usually achieved by the use of "blind" techniques. While physicians have never had any difficulty in accepting the fact that the patient is suggestible, physicians seem to have found it much harder to accept the fact that they themselves are likewise sug-

---

[49] Burk & Buday, *Patient Report Forms: Their Design and Use in Drug Evaluation*, 7 Curr. Ther. Res. 422 (July 1965). See also Sutherland, "The Design of Records and Follow-Up," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 151–154 (1960); Knowelden, *supra*, at 155–159; Truelove, *supra*, at 152; Taber, *supra*, ch. 11; Cox, *supra*, ch. 14–16; Lasagna & Meier, "Experimental Design and Statistical Problems," in Waife & Shapiro, *The Clinical Evaluation of New Drugs* 37, 46 (1959); Modell & Houde, *Factors Influencing Clinical Evaluation of Drugs with Special Reference to the Double Blind Technique*, 167 J. Am. Med. Assn. 2190, 2194 (1958; Carr, *supra*, at 313.

gestible, and that observations would be less than objective if the physician's expectations and biases were not taken into account.[50] There are situations in which the injection of bias of this sort into the experiment due to the preconceptions of the patient and observer is quite easy, and others in which it is highly unlikely. In general, the double blind control should be incorporated into the experiment whenever it is feasible, unless there is very good reason to believe that bias on the part of the patient, observer, and investigator cannot affect the outcome.

> Double-blind and single-blind techniques. The expression "double-blind" (sometimes "double-blindfold") is used to describe a trial in which the nature of the treatment being received by a subject at any time is unknown to both subject and observer. This precaution gives protection against the preconceptions and anticipations of both, and is often required to render the trial valid and the data interpretable.
>
> Provision must always be made to enable the observer to find out immediately what the drug is, if it should be in the patient's interest to do so. The double-blind technique, properly conducted, is in no way unethical.
>
> The term "single-blind" refers to a trial in which one participant, usually the subject rather than the observer or investigator, is unaware of the treatment he is receiving at any specific time. Such a control is almost always less satisfactory than the double-blind technique, and the slight gain in convenience hardly compensates for the potential loss of rigour in the study.[51]

* * * * *

[50] Gold, *supra*, at 41.

[51] World Health Organization, *supra*, at 24.

The person of the physician and his medications have been shown to be symbolic forces of considerable potency in the relief of symptoms, the so-called placebo action. In addition, the hopes of the patient and the therapist, as well as any bias either may have with respect to treatment or experiment, also influence response to treatment. These must be reckoned with in all clinical evaluations.

The patient may want to get better to the extent that he is inclined to see good effects after administration of any new medicament. On the other hand, he may find compensations in his illness and wish to preserve his complaints and hence be inclined to see negative pharmacodynamic effects. "Toxic" effects from placebo are described in the literature, the so-called negative placebo action.

The physician's knowledge of the nature of the medicament is exceedingly important, for, regardless of how much he tries, if he knows the identity of the medicament, he may well communicate this information to the patient. The importance of this unconscious (or conscious) communication has been proved by a study by Batterman and Grossman, in which no difference between aspirin and placebo was detected by the patient unless the physician knew whether he was giving one or the other.

The standard and well-documented procedure is not only to use both placebo and drug which are identical in appearance but also to keep both physician and patient ignorant of what is in use at the time of questioning, examination, and prescription, i.e., the double-blind technique. Both placebo and drug not only should look alike but also should be prescribed and administered alike in order to prevent bias of both patient and physician from producing apparent drug actions. To be a real control, the placebo should also closely imitate the test-drug in any incidental effects such as taste, or other

action perceivable by the patient or physician, but it should be independent of the action under question in the test.

This device, however, does not eliminate bias as an element in the method; it merely attempts to deal with it by equalizing its effects, so that unequally distributed bias alone will not fashion the apparently decisive evidence. At this point, it seems appropriate to emphasize that the double-blind technique is merely a control device; for that reason, we shall designate it the double-blind control.[52]

* * * * *

In the above paragraphs it was remarked that imprecise measurements might make it impossible to detect a substantial actual difference between drugs. An even more serious error, however, is to find and declare statistically significant differences that arise from sources other than true drug effects. Most patients with disease want to get better, and most investigators have some sort of prejudice about any given drug—usually in wanting to come up with successful results, but at times in the opposite direction. This enthusiasm (or lack of enthusiasm) must be allowed to diffuse itself out as equally as possible over the medications under study. The major reliance here must be placed on "blind" technique. Usually, this is "double blind," i.e., patient and observer are both unaware of the nature of a particular medication. At times, a "single blind" technique suffices, if the end point to be determined (such as death) is not particularly amenable to overstatement or understatement, or if the patient records the data himself under circumstances in which the experimenter cannot influence him.

Whether the treatments being compared are active drugs or active drugs and placebos, it is

[52] Model & Houde, *supra*, at 2195.

> necessary (for successful deception) to have tablets or capsules or injections that are as indistinguishable in physical appearance as possible. The medications are then designated by code letters or numbers (preferably a different one for each patient) and the code is known only to certain individuals not directly concerned with the performance of the trial. [Lasagna and Meier (32) p. 47 [53]]

* * * * *

Techniques for avoiding bias in the assessment of therapeutic results have been discussed very fully in the scientific literature.[54]

(9) *The plan or protocol must include a precise statement of the nature of the control group against which the effects of the new treatment modality can be prepared.*

There seems to be universal agreement among those conducting clinical trials that a suitable standard of comparison for the test medication is essential if its therapeutic efficacy is to be properly assessed. The nature of this standard of comparison or "control treatment" will vary depending on the kind of information desired concerning the test medication, the nature of the disease entity being studied, and whether

---

[53] Lasagna & Meier, *Experimental Design & Statistical Problems "in Waife and Shapiro," the Clinical Evaluation of New Drugs* 37, 47 (1959).

[54] Barber, *supra*, at 27; Beecher, *Measurement of Subjective Responses* (1959); Beecher, *Clinical Impression and Clinical Investigation*, 151 J. Am. Med. Assn 44–45 (1953); Goldstein et al., *Principles of Drug Action* 794 (1968); Goldstein, *Biostatistics: An Introductory Text* 9 (1964); Hill, *The Clinical Trial*, 247 N. Eng. J. Med. 113, 117 (1952); Hollister, *Placebology: Sense and Nonsense*, 2 Curr. Ther. Res. 477–483 (1960); Nickerson, *supra*, at 118–122; Schneiderman, *supra*, at 67; Jaber, *supra*, at 86; Truelove, *supra*, at 153; Van Winkle et al., *supra*, at 960; Gold, *supra*, at 41; Carr, *supra*, at 309; Laurence, *supra*, at 12; Gaddum, *supra*, at 197.

standard medications generally accepted as being efficacious in the particular disease already exist.

For quantifying the therapeutic and toxic effects of a new drug, there are two usual standards of reference. One is the placebo, and the other is the drug (or drugs or other forms of therapy) generally accepted as the best treatment already available. The decision whether to include only the placebo, only a standard drug, or both, will depend on the nature of the disease, the drugs already in use for the disease, the state of the relevant experimental methodology, and the goals of the study. Placebos may be crucial to the interpretation of an investigation in which the performance of a new drug appears similar to, or inferior to, the standard medication. Even here, however, the use of placebo controls is not mandatory; the establishment of dose-response curves for the new and the standard drugs may obviate the need for a placebo and will indeed provide a clearer picture of the status of a new drug than a placebo-controlled experiment with single dose levels of new and old drugs.

The placebo is a control for two types of phenomenon. One, the best known and appreciated, is the effect of suggestibility, personality, attitudes, anticipations and other biases on the part of the patient, investigator or observer. These biases may be in the direction of augmenting the benefit of treatment or of diminishing it, of concealing side-effects or of reporting or displaying ill-effects that are unrelated to treatment.

In addition, the placebo provides a vital control for spontaneous changes in the course of the disease or in the symptoms under study, as well as for events that are independent of the treatments under study. In the absence of a period of study with, or of a group of patients receiving, no treatment of any kind (in-

cluding placebo), it is impossible to determine the relative contribution of spontaneous changes in the disease and other events independent of treatment to the overall placebo "success rate" or "toxicity rate." In most instances, however, this distinction will not be important to the investigator, and the placebo will remain an important tool for distinguishing between a true pharmacological effect and both the psychological effects of taking medication and the fortuitous changes associated with the passage of time.[55]

* * * * *

Concurrent or retrospective controls. In between patient comparisons, the patients in the group receiving the treatment under test and those in the control group are generally best allocated concurrently. By doing this, known and unknown fluctuations in the disease, the environment and the type of patient admitted are more likely to be equally represented in each group.

If patients treated in the past are used as controls, the likelihood that the treated and the control groups will be equivalent is much less.

Concurrent controls are particularly important where the natural history of diseases or symptoms is likely to be inconstant and where changes in the kind of patient to be selected, or in the environment, frequently occur. But where the prognosis is fairly constant, retrospective controls may be acceptable and may, indeed, be the only kind that are ethically permissible.[56]

* * * * *

* * * [O]ne of the most important decisions in planning a study is the selection of preparations to be compared. The classical basis of comparison is an inert lactose placebo, which

[55] World Health Organization, *supra*, at 23.

[56] World Health Organization, *supra*, at 24.

allows determination of whether the test substance has any pharmacodynamic effect in the dose administered. This is sometimes a very important and difficult question to answer, but comparison with a completely inert preparation is often far from ideal * * *

* * * [A]t the present stage in the development of drug therapy, the difference between no therapeutic effect and some effect is often not the most important distinction to make. A more important question may be whether the test drug is less effective, equivalent to or more effective than some similar agent of established therapeutic value. . . . Consequently, it is often desirable to select an agent for comparison which ranks as high as possible in proved therapeutic effectiveness and which is as similar as possible to the test drug in its mechanism of action and in the types of side effects produced.[57]

(10) *The plan or protocol must include a survey of statistical methods used in analysis of the data derived from the subjects.*

With most clinical trials, a formal statistical analysis of differences observed is required to minimize the drawing of conclusions unjustified by the magnitude of differences obtained and the number of subjects studied. Appropriate statistical analysis can also greatly increase the amount of information that can

[57] Nickerson, *supra*, at 120. See also Van Winkle et al., *supra*, at 960; Schneiderman, *supra*, at 68; Armitage, "The Construction of Comparable Groups," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 14–18 (1960); Gaddum, *supra*, at 11; Taber, *supra*, at 85; Gold, *supra;* Hill, *Statistical Methods in Clinical and Preventive Medicine* 3–28 (1962); Goldstein et al., *Principles of Drug Action* 794 (1968); Modell & Houde, *supra*, at 2194; Hill, *The Clinical Trial*, 190 Practitioner 85 (January 1963); Doll, *supra*, at 159.

be extracted from a body of data. There are times when the results of a clinical experiment are so clear-cut that a test of statistical significance is unnecessary, but unfortunately, medications are seldom this startlingly effective. The choice of a method of statistical analysis appropriate to a given experimental design and body of data may be quite difficult, and, except in rather obvious cases, the investigator should justify his choice of statistical methodology in the protocol.

If the experiment has been well designed and conducted and has not run into special difficulties, the statistical analysis should create no problems. However, there are a few basic principles that deserve consideration.

First, the nature of the analysis is entirely dependent on the design of the experiment. (This assertion may seem to be a truism, but it is surprising to see how often it is overlooked. For example, one can find many instances in which a sample of paired observations is analyzed as if the observations were independent.) Any experimenter would do well to be sure he knows how to design, make random allocations for, and analyze the simplest kinds of experimental designs—in particular the completely randomized and the randomized-block designs. If, for some reason, a more complex design is needed, he should seek the advice of a competent statistical consultant. It is true that the complex designs and methods for their analysis are carefully described in a number of texts, but the necessary assumptions are not always clearly set forth and, even when they are, the likelihood that the clinical study in question will not materially violate the assumptions is

better decided jointly by the clinician and the statistician than by either alone.[58]

* * * * *

The planning execution and analysis of a therapeutic trial demand statistical ways of thinking and statistical techniques. If a statistician is involved, as consultant or as one of a team, he should preferably be a medical statistician. He should be able to assist in the design of the trial so that comparisons of the various treatments will be made as precisely and as efficiently as possible and with minimal disturbance from other factors that may influence the clinical response.

Often, however, it is not sufficient for the statistician to be involved solely in the original design of the trial and in the final analysis of its results. During the course of a trial many problems may arise, small or great, that demand immediate attention and, possibly, some modification of the original design. These problems may well have statistical implications and the advice of the statistician should be sought as they arise and not after the event, when it is too late to take action.

In the final analysis of the results, one treatment will be contrasted with another. A formal test of statistical significance may then be required. It is important that the clinical investigator understand the meaning of this test. It tells him how often a difference between two (or more) groups of patients of the observed (or greater) magnitude would have arisen by chance. If the difference is unlikely to have arisen by chance, the investigator is in a posi-

[58] Lasagna & Meier, "Experimental Design and Statistical Problems," in Waife & Shapiro, *The Clinical Evaluation of New Drugs* 37, 55 (1959).

> tion to conclude that it is due to the drug he is testing. The better the design and conduct of the trial, the surer he can be of that conclusion.[59]
>
> * * * * *
>
> The fourth major principle refers to the analysis of data in a sophisticated fashion so that one can answer such important questions as "What are the chances that the observed differences between treatments may have been due to chance alone?" or "How reliable are my estimates of the potency of these drugs?" Let it be emphasized that a flagrant disregard of one or more of the principles previously described usually renders it unnecessary to apply any statistical techniques to the data.[60]
>
> * * * * *
>
> It is well to remember that statistics prove nothing—they are merely a device for establishing the betting odds on the reproducibility of the results by mere chance. Statistical prognostication is always based on the assumption that the data used were worthy of collection; statistical analysis of poor data is tantamount to attempting to make a silk purse out of a sow's ear.[61]

A number of excellent texts are devoted to statistics as applied to medical surveys and experiments and include material both on general statistical concepts relating to the design of medical experiments and on specific tests of statistical significance appropriate for various types of controlled clinical trials.[62] There are

---

[59] World Health Organization, *supra*, at 28.

[60] Lasagna, *The Controlled Clinical Trial: Theory and Practice*, 1 J. Chron. Dis. 353, 363 (1955).

[61] Modell & Houde, *supra*, at 2196.

[62] Hill, *Principles of Medical Statistics* (8th ed. 1966; Mainland, *Elementary Medical Statistics* (2d ed. 1963); Goldstein, *Biostatistics: An Introductory Text* (1964).

also symposia and monographs, one or more chapters of which deal with the above points.[63]

---

[63] Armitage, "Statistical Aspects of Clinical Trials," in Zaimis & Elis, *Evaluation of Drugs in Man* 165–170 (1965); Armitage, "The Construction of Comparable Groups," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 14–18 (1960); Beecher, *Measurement of Subjective Responses* (1959); Cox, *supra*; Goldstein et al., *Principles of Drug Action* (1968); Hill, *Statistical Methods in Clinical and Preventive Medicine* 3–28 (1962); Hill, "Aims and Ethics," in Council for International Organizations of Medical Sciences, *Controlled Clinical Trials—A Symposium* 3–7 (1960); Schneiderman, *supra;* Taber, *supra;* Truelove, *supra.*

U.S. GOVERNMENT PRINTING OFFICE: 1973